



BMJ Open is committed to open peer review. As part of this commitment we make the peer review history of every article we publish publicly available.

When an article is published we post the peer reviewers' comments and the authors' responses online. We also post the versions of the paper that were used during peer review. These are the versions that the peer review comments apply to.

The versions of the paper that follow are the versions that were submitted during the peer review process. They are not the versions of record or the final published versions. They should not be cited or distributed as the published version of this manuscript.

BMJ Open is an open access journal and the full, final, typeset and author-corrected version of record of the manuscript is available on our site with no access controls, subscription charges or pay-per-view fees (<http://bmjopen.bmj.com>).

If you have any questions on BMJ Open's open peer review process please email info.bmjopen@bmj.com

BMJ Open

Assessment of the potential for automating administrative tasks in general practice

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-032412
Article Type:	Research
Date Submitted by the Author:	19-Jun-2019
Complete List of Authors:	Willis, Matthew; University of Oxford, Oxford Internet Institute Duckworth, Paul; University of Oxford, Department of Engineering Science Coulter, Angela; University of Oxford, Nuffield Department of Population Health Meyer, Eric; University of Oxford, Oxford Internet Institute; University of Texas at Austin, School of Information Osborne, Michael; University of Oxford, Department of Engineering Science
Keywords:	automation, office, machine learning, PRIMARY CARE, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Title

Assessment of the potential for automating administrative tasks in general practice

Authors

Matthew Willis¹ (Corresponding Author Matthew.willis@oii.ox.ac.uk, +44 01865 287 213); Paul Duckworth²; Angela Coulter³; Eric T. Meyer^{1,4}; Michael Osborne²

1. University of Oxford
Oxford Internet Institute
1 st. Giles
Oxford OX1 3JS
United Kingdom
2. University of Oxford
Department of Engineering Science
Parks Road
Oxford OX1 3PJ
United Kingdom
3. University of Oxford
Nuffield Department of Population Health
Richard Doll Building
Old Road Campus
Oxford OX3 7LF
United Kingdom
4. University of Texas at Austin
School of Information
1616 Guadalupe St
Austin, TX 78701
USA

Keywords

Automation, Office (MeSH Unique ID: D009816); Machine Learning (MeSH Unique ID: D000069550); Primary Health Care (MeSH Unique ID: D011320); Health Policy (MeSH Unique ID: D006291); Organization and Administration (MeSH Unique ID: D009934)

Word Count

4,102

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Objective: To identify the extent to which tasks carried out by primary care staff in general practice could be automated.

Design: A mixed-method design including ethnographic case studies, focus groups, interviews and an online survey of automation experts.

Setting: Three urban and three rural general practice health centres in England selected for differences in list size and organisational characteristics.

Participants: Observation and interviews with 65 primary care staff in the following job roles: administrator, manager, general practitioner, healthcare assistant, nurse practitioner, pharmacy technician, phlebotomist, practice nurse, pharmacist, prescription clerk, receptionist, secretary, and summarizer; together with a survey of 156 experts in automation technologies.

Methods: 330 hours of ethnographic observation and documentation of administrative tasks carried out by staff in each of the above job roles, followed by coding and classification; semi-structured interviews with 10 GPs, and six staff focus groups. The online survey of machine learning, artificial intelligence, and robotics experts was analysed using an ordinal Gaussian process prediction model to estimate the automatability of the observed tasks.

Results: The model predicted that roughly 44% of administrative tasks carried out by staff in general practice are “mostly” or “completely” automatable using currently available technology. Discussions with practice staff underlined the need for a cautious approach to implementation.

Conclusions: There is considerable potential to extend the use of automation in primary care, but this will require careful implementation and ongoing evaluation.

Strengths and limitations

Strengths

- This is currently the first use of a scalable measure of task-based automation applied to the work of healthcare. Specifically, it is grounded in empirical data from NHS England primary care work practices.
- Using an Independent Bayesian Classifier Combination to achieve a single rating from multiple experts is advantageous both because it is fully Bayesian and reflects a higher

chance of accurately recovering the true automatability label of a task in an environment of uncertainty and subjectivity.

Limitations

- This study was only able to analyse administrative tasks, our use of the O*NET database in large part restricted us to looking at these administrative functions.
- Primary care looks different across all parts of the country and unfortunately, we were unable to gather data on some of these nuances. We made extensive efforts to understand the tasks performed by each occupation and to ensure they were as representative as possible, but we cannot be certain that we captured all the diversity.
- The expert survey could have captured data about how a task could be automated, for example, by simplifying or re-arranging the task. This subjectiveness in the task-simplification may have translated into our dataset providing very few entirely “not automatable tasks. Caused, in part, by the creativity of the responder’s who would believe part of a re-arranged task to be automatable.

Introduction

Primary care in England faces numerous challenges, including increased demand, workload pressures, recruitment difficulties, and budgetary constraints [1–4]. All staff in general practice, including GPs, nurses, pharmacists, and non-clinical staff, carry out many administrative tasks. There has been much discussion of the bureaucratic burden on general practice and how to streamline these tasks to free up more time for direct patient care. A 2015 report identified several potential areas of concern, including time spent on activities that concern ‘getting paid’, entering data into various unintegrated practice-based information systems, processing information from hospitals and other external providers, keeping on top of changing requirements of commissioners, and helping patients navigate a fragmented and poorly coordinated system [5]. The authors concluded that over 50% of practice time was spent on bureaucracy and that much of this was potentially avoidable. A majority of GPs nowadays describe their workload as excessive and detrimental to the delivery of high-quality care [6–8]. This phenomenon of ‘paperwork’ growth is consistent with the general upward trend in bureaucratic work in public sector occupations [9]. In general practice in England, administrative, non-clinical staff now outnumber clinical staff by a factor of 2.5 to 1 [10]. The Royal College of General Practitioners has recently called for significant investment in digital technology to transform care, including automation of routine administrative tasks [11].

Automation of administrative tasks may reduce the bureaucratic burden in primary care, but the nature and extent of its likely impact is currently unknown. In other sectors automation is typically perceived as a threat to workers, conjuring visions of mass unemployment. Previous research by

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

members of our team has classified 47% of occupations as being “high risk” for computer automation within the next couple of decades [12]. That study based its estimates of the probability of automatability on the US Department of Labor’s Occupational Information Network (O*NET) database of occupational characteristics and worker requirements [13]. Later studies applied the same probability estimates to employment data from other countries, on the assumption that the probabilities are consistent globally. The proportion of jobs estimated as ‘high risk’ ranged from 45% to more than 60% across different European countries [14–16]. However, where there are workforce shortages as in general practice, automation presents not a threat but an opportunity, potentially increasing productivity and leading to more satisfying job roles. The assumption has been that healthcare occupations are less likely to be affected because they rely heavily on interpersonal interactions, but they also include tasks that require information acquisition, information analysis, and decision making, all characteristics amenable to automation[17]. In primary care, automatable tasks include checking written documents for errors, reviewing and writing letters, creating referral letters, organising staff rotas, creating reports and maintaining records. Technologies such as natural language processing, voice recognition, text summarisation, robotic process automation, and machine learning to support the manipulation of information are all potentially applicable [18,19]. It is unlikely that implementation of automation technologies in general practice will cause large-scale unemployment of healthcare staff, but it could enhance their capacity to keep on top of administrative work while freeing up more time for patient care.

Our aim in carrying out the present study was to quantify the likely impact of automation in general practice to inform the assessment of future resource needs. This was not a future-gazing project. We were interested in the potential for automation using currently available technologies, rather than those that might become available at some point in the future.

Methods

This study deployed mixed qualitative and quantitative data collection and analysis. To help us make this assessment, we carried out observations in six general practices, conducted a survey of experts on automation technologies and used a predictive machine learning model to apply their assessments to primary care tasks, grouped according to the O*NET classification system.

Qualitative data collection and analysis

The first stage involved 330 hours of ethnographic observations by author MW across six primary care centres in England located in Oxfordshire, Yorkshire, Berkshire, Surrey, and the West Midlands. The practices' list sizes ranged from 5,000 to 24,000 patients. The researcher spent one week in each location, observing work carried out by GPs, other primary care clinicians and non-clinical staff. The focus of the fieldwork was to describe and classify tasks carried out by each of the following occupational groups: *Administrator, Deputy Practice Manager, General Practitioner, Healthcare Assistant, Nurse Practitioner, Pharmacy Technician, Phlebotomist, Practice Manager, Practice Nurse, Practice Pharmacist, Prescription Clerk, Receptionist, Secretary, and Summarizer*. Data collected during the fieldwork consisted of notes, photographs and documents that elucidate tasks being observed. A focus group was held on the final day at each site to gather staff feedback on the researcher's observations and to check that tasks observed were representative of their occupations. A brief survey describing 15 random tasks per occupation was later sent to one of the practices for further validation. Author MW developed a spreadsheet of each observed occupation and a list of the tasks they perform. Task data were coded in the spreadsheet and further enriched through a series of categories that include the frequency of task, any tools or technologies used, type of software required, and other features of the task including conceptual inductive and deductive categories. Important themes were elicited and refined after a continuous process of conversation and collaborative sensemaking between the authors. All fieldwork involving interactions with staff was approved by the Health Research Authority under project number 212952.

Quantitative data collection and analysis

The O*NET system describes occupations in terms of the knowledge, skills, and abilities required as well as how the work is performed in terms of tasks, work activities, and other descriptors. It provides a useful basis for an analysis of automatability, but it was not validated for use with British healthcare-specific occupation and task data. Our first task, therefore, was to carry out observation of primary care tasks to enable linkage to the O*NET system. The focus on task data is key because it is individual tasks rather than entire occupations that are typically automated[20–24] (Bresnahan et al., 2002; Brynjolfsson and McAfee, 2012, 2014; Duckworth et al., 2018; Susskind and Susskind, 2017). Recent research suggests that occupations are best analysed as evolving combinations of detailed tasks, skills, and/or environments [23,25–27]. With more granular job data available than ever before, thousands of occupations can each be broken down into hundreds of numeric components or tasks, relating to the skills, knowledge, and abilities required to perform them.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

The method conventionally used to identify and predict the likelihood of automation is to hypothesize the underlying dynamics that drive automation and extrapolate into the future an unspecified number of years [12,25,28–30]. This relies on subjective forecasts that may be unreliable. Instead, we decided to ask experts to base their estimates on *current* technology capabilities, rather than speculate on the future. To obtain current estimates of the automation potential of tasks, we contacted experts in the machine learning, robotics and artificial intelligence communities through academic mailing lists and lists acquired at two machine learning conferences (specifically the Conference on Neural Information Processing Systems, 2015 and 2017). Experts verified their academic and industrial experience and (optionally) their contact information.

Each survey participant was sequentially presented with five O*NET occupations and the five tasks with the largest task-importance score for that occupation. Task importance scores are rankings of their importance to the occupation that performs the task and are provided by O*NET. Participants were asked the following question: *“Do you believe that technology exists today that could automate these tasks?”*

Each task was ranked on an ordinal scale from 1.0 to 4.0: not automatable today (1.0), mostly not automatable today (human does most of it) (2.0), could be mostly automated today (human still needed) (3.0), completely automatable today. (4.0), or unsure. Respondents were also asked to rate their overall confidence in their answers. In total, we received 4,599 task level responses from 156 academic and industrial experts around the world, shown in Figure 1 and 2.

The multiple ratings or labels for each task were combined with Independent Bayesian Classifier Combination (IBCC), a principled Bayesian approach to combine multiple classifications [31,32]. IBCC creates a posterior prediction over labels that reflects the tendency of the individual labellers to concur with ultimately chosen label values. IBCC task scores were averaged into 313 work activities (note: an activity is a unique set of tasks in O*NET taxonomy). Therefore, activity automatability labels concentrate around whole and half values. The final values were rounded to the closest 0.5 (a half-class). A score of 4.0 indicates a “fully automatable” work activity, while 1.0 indicates a work activity that “cannot be automated” using currently available technology [23].

Finally, all data sources were combined to form the basis of the analysis, including: 1) the validated list of primary care tasks obtained from the ethnographic observations, interviews and focus groups, coded and classified allowing us to use 2) the corresponding O*NET activity skills, knowledge, and ability numerical attributes, and 3) the inferred automatability scores of these activities informed by

the expert survey. A machine learning framework was employed to infer a functional mapping between the skills, knowledge, and ability characteristics of work activities and the ground truth automatability elicited from the expert survey. The anticipated output was a flexible function estimation capable of modelling complex, non-linear relationships between the features (skills, abilities, knowledge) and the target (automatability).

The collection of work activities (N=313) labelled in the expert survey of automatability was used as a training data set for the Gaussian Process Ordinal Regression model. Once trained, the model can estimate the automatability of “unlabelled” work activities, i.e. our observed healthcare tasks. This is preferable to asking experts to provide judgements on the full range of tasks observed, which would be time-consuming and prohibitively expensive. A Kernel Density Estimation (KDE) analysis was employed to show the distribution of the inferred automatability of the healthcare tasks. KDE allows visualization of a smooth curve that represents the “shape” of the data distribution as a replacement for a discrete histogram. Further technical details of the quantitative framework and analysis, as well as a detailed description of the observational methods, are documented in a previous publication [33].

Patient and Public Involvement

No patients were involved in this study.

Results

Administrative tasks

The ethnographic study resulted in 65 unique sessions of observation of work carried out by primary care staff, each ranging from 45 minutes to 90 minutes. These observations can be best illustrated by focusing on two large and significant types of work that account for many of the administrative tasks carried out – letter work and clinical documentation. It is important to note that although we classify this as administrative work, we appreciate the medical importance of the tasks and that they involve several sub-tasks, some of which require specialist medical knowledge.

Letters

This is the bundle of tasks and workflows that relate to the production of letters. Letter production flows in two directions; incoming and outgoing. Incoming letters contain multiple forms of implicit and explicit information. Processing this information results in tasks such as responding to requests, storing physical and digital copies of the letter, applying a medical coding scheme, and using information contained in the letter to inform or modify patient treatment and/or other clinical decisions. Outgoing letters involve multiple processes as well, such as audio dictation, transcription

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

of audio text, application of preformatted templates, appending directions and notes to audio files or digital documents, referencing information in the electronic medical record, archiving letters, and posting the letter. The specific details of how this work is carried out varies from practice to practice. The workflow can look remarkably different in a single-site independent practice when compared to the workflow in a super-practice, sharing services across multiple sites. A single-site practice may allocate this work to receptionists who process letters, while a super-practice might outsource all medical coding, audio transcription, and typing to a central location.

The following workflow was observed at a small single site practice. Receptionists sort incoming letters into groups, giving them red, amber, or green status dependent on response urgency. Then, they scan the letters into the system and use the commercial software system DocMan to sort them using the colour-coded system. One of the receptionists is assigned the task of “dealing” with the letters. This involves opening a letter that has been briefly coded by a GP with the highlighter tool in the document management software, reviewing the highlighted medical terms or phrases, and allocating appropriate Read codes. The letter is then archived on the system, the task is marked complete, and the receptionist continues to the next letter. If the receptionist has a question about the appropriate Read code, text comments are added to the document, which is then saved, and assigned as a task on the system, i.e. sent back to the GP. Other types of letters, for example, from a medical supplier requesting information about a patient who requires a specific device, must find the required information in the relevant health record, then use a template letter to respond to the supplier.

The length of letters ranges from precise and short to long and verbose. Letters have a different utility dependent on who is writing them, who is reading them, and the context. For example, during our research a GP was observed reading a 9-page letter from secondary care. The GP skim-read the letter on her monitor and after a few minutes declared, “ah hah, you see here it is”, highlighting a paragraph with the mouse cursor, saying “this is all I need, out of this whole document, this is what is relevant to me.” She continues: “I can tell the rest of this letter is for their [the specialist’s office] purposes, it’s a form of documentation, I don’t need all this.” The GP did not require most of the content of the letter, only a single paragraph. Her prior experience of the specialist who wrote the letter enabled her to extract the relevant information efficiently and discount the rest.

Clinical documentation

A scenario in which a system could automate clinical documentation using natural language processing and text summarisation was presented in focus groups and interviews, provoking lengthy

discussion. Some participants saw it as a potentially useful technology, with multiple “when can we have that” comments, while others felt it might have negative consequences on the clinical encounter and could even undermine the skills and critical thinking of the clinician. One GP argued that having a transcript of the entire conversation would be a waste of time, producing a great deal of unnecessary information. It is not uncommon for GPs to re-write or edit their notes, particularly notes for new cases or complex circumstances, or when the notes are written by trainees. What may look like a relatively simple administrative task in fact involves a considerable amount of clinical skill and experience. Removing this task from the clinical workflow would remove an opportunity for the clinician to reflect critically about an interaction. Research shows that writing engages the brain, allows GPs to be better observers, supports empathy, and engages critical thinking [34].

Scope for automation

The online survey of experts in automation systems achieved 156 responses. The distribution of respondents by country and qualification is shown in figures 1 and 2.

[Figure 1: Number of survey respondents by country.]

[Figure 2: Number of survey respondents by qualification.]

The Gaussian Process machine learning model that we have applied to the data suggests that about 44% of the administrative tasks performed in primary care are “mostly” or “completely” automatable using current technology. Figure 3 shows the distribution of automatability for all administrative tasks performed by general practice staff.

[Figure 3: Distribution of Automatability Scores. Plot shows the Kernel Density Estimate of the Scores.]

As discussed above, tasks performed by various occupations overlap, but the automatability of those tasks are independent of the occupation, i.e. an automatability score is not dependent on who performs the task. For example, medical coding has an estimated automatability of 3.2. Thus, automation of this task, which is mostly performed by receptionists, summarisers, and secretaries, but also on occasion by GPs depending on the practice, could either remove this task from their

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

workloads altogether, if the systems are considered sufficiently reliable, or reduce it to checking codes produced by an automated system. Some of the most highly automatable tasks in our data include: *payroll and managing finances, checking and sorting post, printing letters, communicating with patients through texting, management of paper archives (onsite or offsite), transcription, email account management, letter scanning, checking for errors in paperwork, and internal communications (for example messages to staff or new employee inductions).*

To further understand the automatability predictions, three categories were generated: automatable (predicted score is >3.0), not-automatable (predicted score < 2.0), and partly-automatable (predicted score ≥2.0 to ≤3.0). Figure 3 presents the predicted automatability scores of administrative occupation tasks classified into these risk categories.

[Figure 4: Percentage of Administrative Tasks performed in Primary Care by Automatability Categories.]

Scanning clerks are typically, but not always, part time employees that are responsible for scanning physical documents (mostly letters) and creating digital objects (usually a PDF file) and attach these files to the patients’ health record or import files into other document management systems. Scanning clerks also add meta data to the PDF so the files are searchable and contain all contextual information. Results show that only the work of scanning clerks could be fully automated using currently available technology, thus eliminating the need for that occupation entirely. Many routine administrative tasks and typical office work (e.g. answering phones, writing letters and managing email) are undertaken by all occupational groups, but the impact of automation will vary according to the number of tasks in each automatable category that they perform. Figures 4 and 5 below show the potential impact of automation technologies by occupation for clinical and non-clinical staff.

[Figure 5: (KDE) Distribution of Task Automatability for Primary Care Occupations (non-clinical staff).]

[Figure 6: (KDE) Distribution of Task Automatability for Primary Care Occupations (clinical staff).]

Further detailed results can be found in the technical report for The Health foundation, currently in preparation [35].

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Discussion

We believe this study is the first to combine detailed validated observations of workflows and a specially developed ordinal scale of automatability to assess the potential for automating various administrative tasks routinely carried out by primary care staff. Our findings indicate considerable scope for extending the use of automation in general practice. It is important to acknowledge that certain capabilities of automation already exist in primary care. In some practices scanning clerks put paper documents into scanners that turn physical text into digital type using optical character recognition. Another form of automation is the use of templates and pre-generated text in desktop word processors or in electronic medical record notes. Automated report generation and custom search queries to input patient data to different systems is another form of automation that is sometimes used. Our analysis has looked at the scope for further automation of tasks and we conclude that it is considerable.

A few caveats must be borne in mind. Observations took place in only six general practices. These varied in terms of list size and geographical location, but we cannot claim that these were representative of all practices in England. The survey of automation experts was, we believe, the largest of its kind to date and it provides a robust baseline estimate of the extent to which tasks are automatable using only currently available technology, but participants did not have specific expertise in the application of automated systems to healthcare. We made up for this deficiency by inviting discussion and feedback from primary care staff involved in interviews and focus groups, but they, in turn, lacked direct experience of using the automated systems that might eventually modify their working practices.

Automation technologies will be most useful when they provide clinicians with the right information at the right time; employing filters of context and urgency to remove details that are irrelevant or distracting. Automation of some mundane tasks, such as telephone answering, letter writing, document scanning, email monitoring, and information filtering, could free up valuable time for more rewarding tasks, such as direct patient care, leading to improvements in workflows and staff satisfaction. We find that many tasks involving data entry, sending information, formatting information, maintaining records, transcription, and operating office equipment to name a few, have high potential for automation. Many practices have basic forms of automated telephone answering, but more sophisticated systems are now available that would enable a voice assistant to help triage the hundreds of phone calls that a practice can receive in a day. However, there are drawbacks to automated phone systems that would need careful consideration and evaluation. Receptionists help prevent unneeded appointments by screening phone calls and getting patients to the right area for

their appointment, these skills rely on social perception and an understanding of the patient's history. Automated phone systems must be aware of this function and help guide patients through their case without scheduling unneeded appointments.

Clinical documentation could benefit from automation for three main reasons. First, these are tasks that every clinician must complete with a degree of urgency. It is considered good medical practice to document the patient consultation as soon as possible, i.e. it cannot be left to the end of the day. Second, documentation requires attention to detail and is time intensive. Even very brief encounters require accurate documentation. Third, this work is typically squeezed into a busy schedule, usually performed during the time it takes a patient to leave the examination room and the next patient arriving. Vendors of electronic health record systems have advocated the use of automation to facilitate record-keeping by use of structured data formats [36–38]. Natural language processing can be used to dissect patient-doctor conversations and create notes [39]. How to ensure the validity and reliability of such systems has been the subject of much research and remains a controversial topic [18,40]. However, a recent survey of UK GPs revealed considerable optimism about the likelihood that clinical documentation would be fully automated in the near future, with a large majority believing this would occur within ten years [41].

Automation technologies can be expensive and time-consuming to implement, so those contemplating the purchase of such systems should consider the following issues: Which types of tasks would be most beneficial to the practice if automated? Are those tasks amenable to automation? If this results in a change in occupational roles, who will it affect most? How will staff be trained and supported to use the new technology? If tasks are removed from the workload of a staff member, what kind of work (if any) will replace the task? And who will be responsible for the maintenance of the tasks while new workflows are in development?

Some tasks with high automatability scores will be challenging to implement for social and organisational reasons. Previous research has shown that staff contribute to patient quality and safety through hidden ways that emerge in the moment and require tacit knowledge [42]. The tacit and contextual work, such as reading social cues, currently lies outside the capabilities of intelligent systems [29]. Healthcare is crucially dependent on these human skills and patient care would suffer if automation is pursued to their detriment. Our observations revealed a type of work that is often driven by exceptions, interruptions, competing critical requests and time sensitivity. Work practices that may appear routine, inefficient or superfluous, often have an intrinsic value for cooperative teamwork [43]. While our quantitative findings reveal considerable potential for automation, our in-depth fieldwork demonstrated the need for a cautious approach.

Conclusion

We believe that careful introduction of currently available technologies has considerable potential to reform administrative workflows and increase productivity in primary care. However, this should not be seen as a magic bullet that will solve the workload crisis in the immediate future. Introduction of new forms of automation should proceed cautiously with ongoing evaluation of their wider impacts.

Competing interests

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: all authors had financial support from The Health Foundation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Author Contributions

Matthew Willis: Writing, coordination of authors, literature search, citation management, study design, formatting, draft iteration, data collection, data interpretation, manuscript submission.

Paul Duckworth: Figures, draft writing, editing, data collection, data analysis, literature search, data interpretation.

Angela Coulter: Reviewing and editing, writing, data interpretation, study design.

Eric T. Meyer: Reviewing and editing, writing, data interpretation, study design.

Michael Osborne: Data analysis, data interpretation, technical advising, figures, data collection, study design.

Funding

This work was supported by The Health Foundation, Award #7559. The study funders played no role in the study design, writing of the report, or the decision to submit the manuscript for publication. Researchers were independent of influence from study funders.

Data Sharing

All data relevant to the study are included in the article.

References

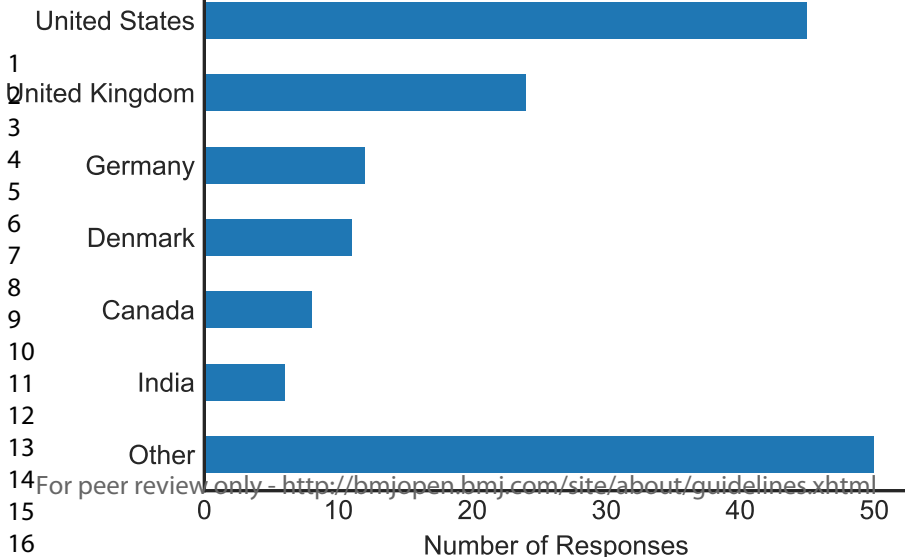
- 1 Hopson C. The State of the NHS provider sector. 2016. <https://www.nhsproviders.org/news-blogs/news/nhs-trust-leaders-warn-staff-shortages-now-outweigh-fears-over-funding>
- 2 Baird B, Charles A, Honeyman M, *et al*. Understanding pressures in general practice. London, UK: 2016. <https://www.kingsfund.org.uk/publications/pressures-in-general-practice>
- 3 Roland M, Everington S. Tackling the crisis in general practice. *BMJ* 2016;**352**:i942. doi:10.1136/bmj.i942

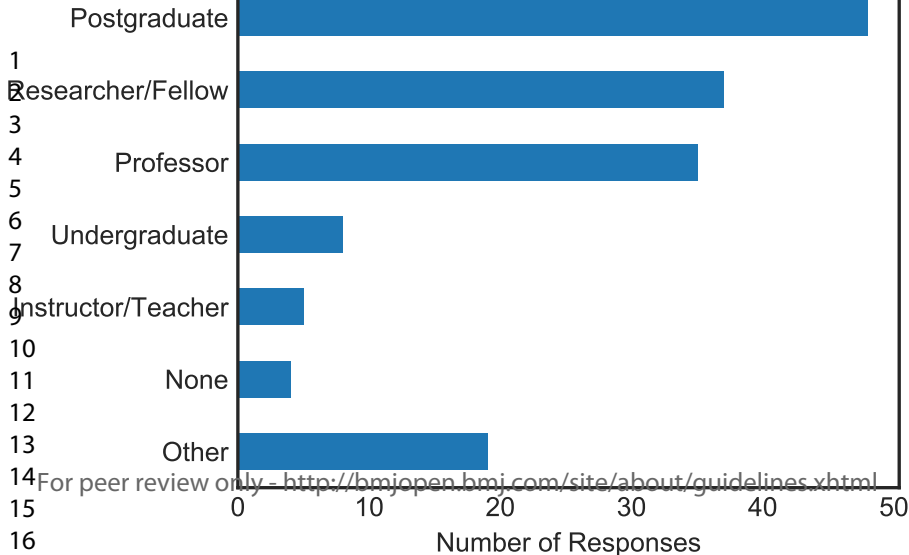
- 4 General Practice Forward View. NHS Engl. 2016.<https://www.england.nhs.uk/gp/gpfv/about/>
- 5 Clay H, Stern R. Making Time in General Practice. Prim. Care Found. Alliance. 2015;;1–83.
- 6 Byrne L, Bottomley J, Turk A. British Medical Association Survey of GPs in England. London: 2016.
- 7 Gibson J, Checkland K, Coleman A, *et al.* Eighth National GP Worklife Survey. Manchester: 2016.
<http://research.bmh.manchester.ac.uk/healthconomics/research/Reports/EighthNationalGPWorklifeSurveyreport/>
- 8 Hobbs FDR, Bankhead C, Mukhtar T, *et al.* Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007–14. *Lancet (London, England)* 2016;**387**:2323–30. doi:10.1016/S0140-6736(16)00620-6
- 9 Hitchcock ALK, Sundorph E. Work in Progress: Towards a leaner, smarter public-sector workforce. 2017.
- 10 NHS statistics, facts and figures. London, UK: 2017.
<https://www.nhsconfed.org/resources/key-statistics-on-the-nhs>
- 11 Fit for the Future: a Vision for General Practice. London, UK: 2019.
<https://www.rcgp.org.uk/policy/future-vision.aspx>
- 12 Frey CB, Osborne MA. The future of employment: How susceptible are jobs to computerisation? *Technol Forecast Soc Change* 2017;**114**:254–80. doi:10.1016/j.techfore.2016.08.019
- 13 National Center for O*NET Development. 2016.https://www.doleta.gov/programs/onet/eta_default.cfm
- 14 Jeremy Bowles. The computerisation of european jobs. Bruegel. 2014.<http://bruegel.org/2014/07/the-computerisation-of-european-jobs/>
- 15 Pajarinen M, Rouvinen P, Pajarinen M, *et al.* Computerization Threatens One Third of Finnish Employment. Published Online First: 13 January 2014.<https://econpapers.repec.org/paper/rifbriefs/22.htm> (accessed 27 May 2019).
- 16 Knowles-Cutler A, Frey C, Osborne M. Agile Town: The Relentless March of Technology and London's Response. London, UK: 2014.
<https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/uk-futures/london-futures-agiletown.pdf>
- 17 Jamieson GA, Skraaning G. Levels of Automation in Human Factors Models for Automation Design: Why We Might Consider Throwing the Baby Out With the Bathwater. *J Cogn Eng Decis Mak* 2018;**12**:42–9. doi:10.1177/1555343417732856
- 18 Kreimeyer K, Foster M, Pandey A, *et al.* Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review. *J Biomed Inform* 2017;**73**:14–29. doi:10.1016/j.jbi.2017.07.012
- 19 Ng A. What Artificial Intelligence Can and Cant Do Right Now. *Harv Bus Rev*<https://hbr.org/2016/11/%0Awhat-artificial-intelligence-can-and-cant-do-right-now>
- 20 Bresnahan T, Brynjolfsson E, Hitt L. Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-Level Evidence. Cambridge, MA: 1999. doi:10.3386/w7136

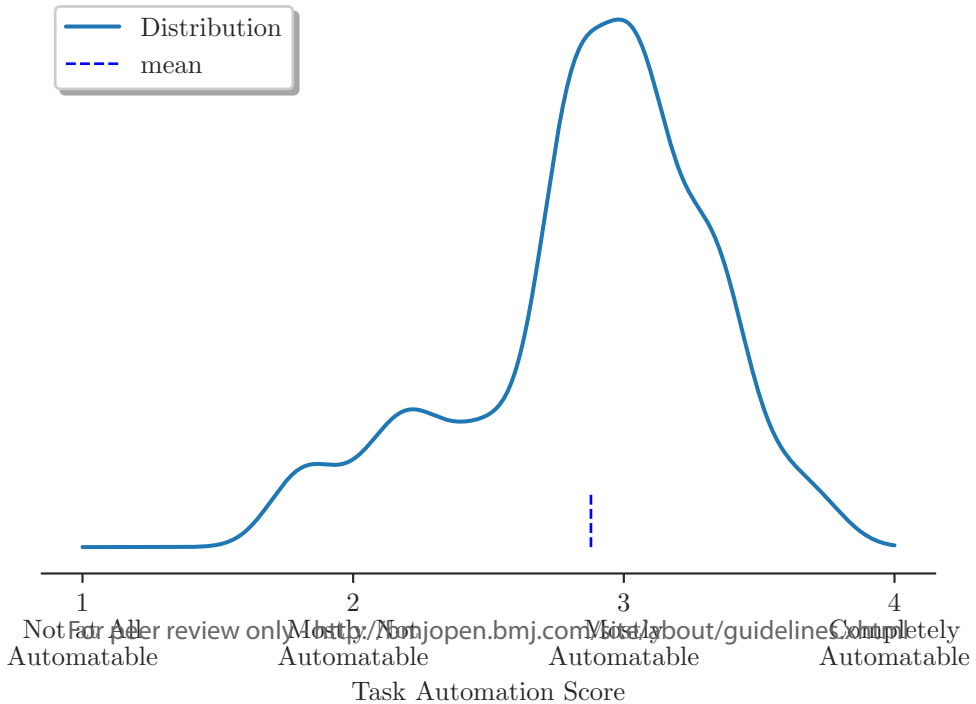
- 21 Brynjolfsson E, McAfee A. *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*. 2012.
- 22 Brynjolfsson E, McAfee A. *The second machine age : work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company 2014.
- 23 Duckworth P, Graham L, Osborne M. Inferring Work Task Automatability from AI Expert Evidence. In: *ACM conference on Artificial Intelligence, Ethics, and Society*. Honolulu, Hawaii: 2019. <http://www.aies-conference.com/accepted-papers/>
- 24 Susskind RE, Susskind D. *The future of the professions: how technology will transform the work of human experts*. OUP Oxford 2017.
- 25 Acemoglu D, Restrepo P. Robots and Jobs: Evidence From Us Labor Markets. *Nber Work Pap Ser* 2017.
- 26 Arntz M, Gregory T, Zierahn U. Revisiting the risk of automation. *Econ Lett* 2017;**159**:157–60. doi:10.1016/j.econlet.2017.07.001
- 27 James Manyika, Michael Chui, Mehdi Miremadi, *et al*. Harnessing automation for a future that works. McKinsey Glob. Inst. 2017.<http://www.mckinsey.com/global-themes/digital-disruption/harnessing-automation-for-a-future-that-works>
- 28 Autor DH, Handel MJ. Putting Tasks to the Test : Human Capital, Job Tasks, and Wages. *J Labor Econ* 2013;**31**:59–96. doi:10.1086/669332
- 29 Bakhshi H, Downing JM, Osborne MA, *et al*. The Future of Skills: Employment in 2030. 2017.
- 30 Goos M, Manning A, Salomons A. Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *Am Econ Rev* 2014;**104**:2509–26. doi:10.1257/aer.104.8.2509
- 31 Kim H-C, Ghahramani Z. Bayesian Classifier Combination. In: *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*. PMLR 2012. 619–627.<http://proceedings.mlr.press/v22/kim12.html>
- 32 Simpson E, Roberts S, Psorakis I, *et al*. Dynamic Bayesian Combination of Multiple Imperfect Classifiers. Published Online First: 8 June 2012.<http://arxiv.org/abs/1206.1831>
- 33 Willis M, Duckworth P, Coulter A, *et al*. The Future of Health Care: Protocol for Measuring the Potential of Task Automation Grounded in the National Health Service Primary Care System. *JMIR Res Protoc* 2019;**8**:e11232. doi:10.2196/11232
- 34 Wald HS, Borkan JM, Taylor JS, *et al*. Fostering and Evaluating Reflective Capacity in Medical Education: Developing the REFLECT Rubric for Assessing Reflective Writing. *Acad Med* 2012;**87**:41–50. doi:10.1097/ACM.0b013e31823b55fa
- 35 Willis M, Duckworth P, Graham L, *et al*. Future of Healthcare: Computerisation, Automation, and General Practice Services. London, UK: 2019.
- 36 Rule A, Rick S, Chiu M, *et al*. Validating free-text order entry for a note-centric EHR. *AMIA . Annu Symp proceedings AMIA Symp* 2015;**2015**:1103–10.<http://www.ncbi.nlm.nih.gov/pubmed/26958249> (accessed 9 Sep 2018).
- 37 Lin SY, Shanafelt TD, Asch SM. Reimagining Clinical Documentation With Artificial Intelligence. *Mayo Clin Proc* 2018;**93**:563–5. doi:10.1016/j.mayocp.2018.02.016
- 38 Verghese A, Shah NH, Harrington RA. What This Computer Needs Is a Physician: Humanism

- and Artificial Intelligence. *JAMA* 2018;**319**:19. doi:10.1001/jama.2017.19198
- 39 Klann JG, Szolovits P. An intelligent listening framework for capturing encounter notes from a doctor-patient dialog. *BMC Med Inform Decis Mak* 2009;**9 Suppl 1**:S3. doi:10.1186/1472-6947-9-S1-S3
- 40 Bansler JP, Havn EC, Schmidt K, *et al*. Cooperative Epistemic Work in Medical Practice: An Analysis of Physicians' Clinical Notes. *Comput Support Coop Work* 2016;**25**:503–46. doi:10.1007/s10606-016-9261-x
- 41 Blease C, Bernstein MH, Gaab J, *et al*. Computerization and the future of primary care: A survey of general practitioners in the UK. *PLoS* 2018;**13**:1–12. doi:10.1371/journal.pone.0207418 December
- 42 Swinglehurst D, Greenhalgh T, Russell J, *et al*. Receptionist input to quality and safety in repeat prescribing in UK general practice: ethnographic case study. *BMJ* 2011;**343**:d6788. doi:10.1136/bmj.d6788
- 43 Cabitza F, Ellingsen G, Locoro A, *et al*. Repetita luvant : Exploring and Supporting Redundancy in Hospital Practices. *Comput Support Coop Work* 2018.

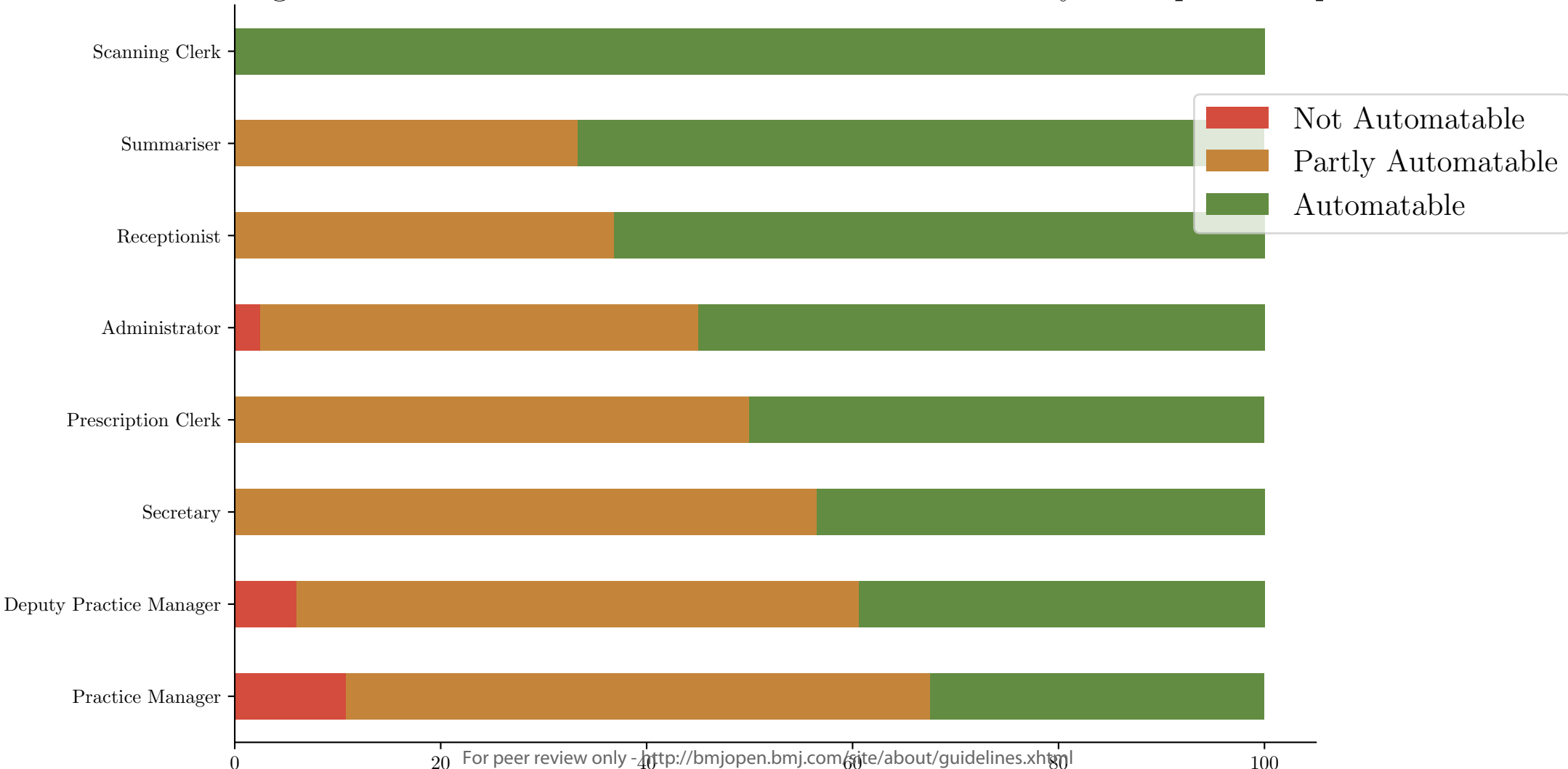
Top Responder Locations



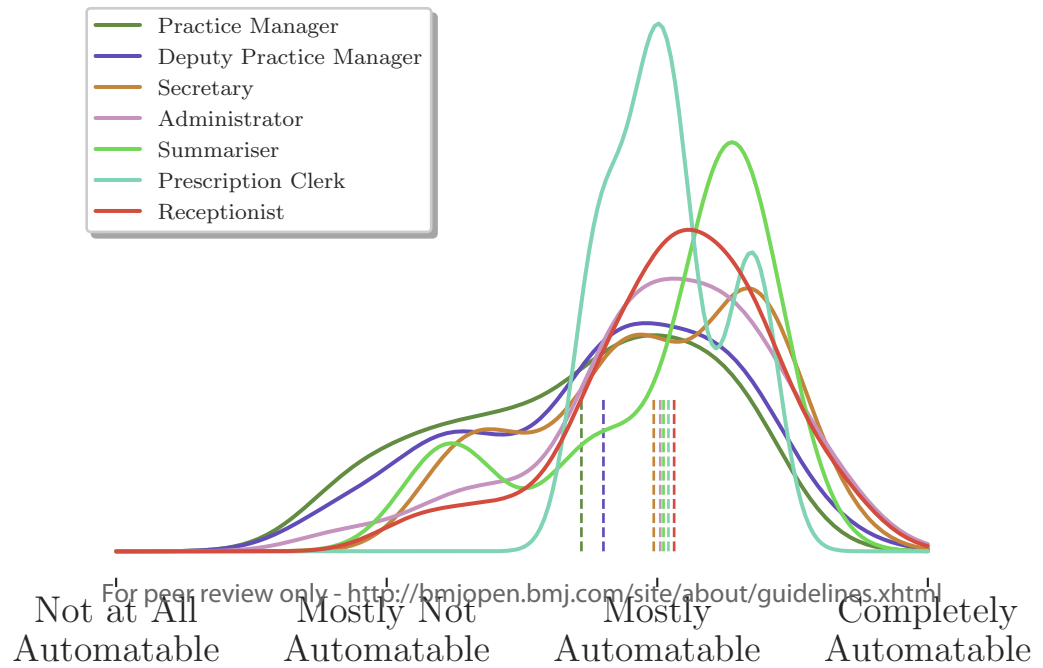




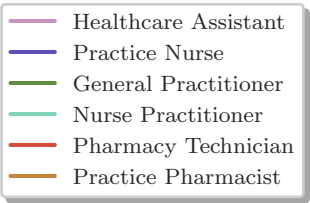
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34



Percentage of Tasks Performed by Automatable Category



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20



Not at All Automatable Mostly Not Automatable Mostly Automatable Completely Automatable

For peer review only - <http://bmjopen.bmj.com/site/about/guidelines.xhtml>

BMJ Open

A qualitative and quantitative approach to assess of the potential for automating administrative tasks in general practice

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-032412.R1
Article Type:	Original research
Date Submitted by the Author:	21-Jan-2020
Complete List of Authors:	Willis, Matthew; University of Oxford, Oxford Internet Institute; University of Michigan, School of Information Duckworth, Paul; University of Oxford, Department of Engineering Science Coulter, Angela; University of Oxford, Nuffield Department of Population Health Meyer, Eric; University of Oxford, Oxford Internet Institute; University of Texas at Austin, School of Information Osborne, Michael; University of Oxford, Department of Engineering Science
Primary Subject Heading:	General practice / Family practice
Secondary Subject Heading:	Health informatics
Keywords:	automation, office, machine learning, PRIMARY CARE, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Title

A qualitative and quantitative approach to assess of the potential for automating administrative tasks in general practice

Authors

Matthew Willis^{1,5} (Corresponding Author Matthew.willis@oii.ox.ac.uk, +44 01865 287 213); Paul Duckworth²; Angela Coulter³; Eric T. Meyer^{1,4}; Michael Osborne²

1. University of Oxford
Oxford Internet Institute
1 st. Giles
Oxford OX1 3JS
United Kingdom
2. University of Oxford
Department of Engineering Science
Parks Road
Oxford OX1 3PJ
United Kingdom
3. University of Oxford
Nuffield Department of Population Health
Richard Doll Building
Old Road Campus
Oxford OX3 7LF
United Kingdom
4. University of Texas at Austin
School of Information
1616 Guadalupe St
Austin, TX 78701
USA
5. University of Michigan at Ann Arbor
School of Information
105 S State St
Ann Arbor, MI 48109
USA

Keywords

Automation, Office (MeSH Unique ID: D009816); Machine Learning (MeSH Unique ID: D000069550); Primary Health Care (MeSH Unique ID: D011320); Health Policy (MeSH Unique ID: D006291); Organization and Administration (MeSH Unique ID: D009934)

Word Count

5,180

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Objective: To identify the extent to which administrative tasks carried out by primary care staff in general practice could be automated.

Design: A mixed-method design including ethnographic case studies, focus groups, interviews and an online survey of automation experts.

Setting: Three urban and three rural general practice health centres in England selected for differences in list size and organisational characteristics.

Participants: Observation and interviews with 65 primary care staff in the following job roles: administrator, manager, general practitioner, healthcare assistant, nurse practitioner, pharmacy technician, phlebotomist, practice nurse, pharmacist, prescription clerk, receptionist, scanning clerk, secretary, and medical summariser; together with a survey of 156 experts in automation technologies.

Methods: 330 hours of ethnographic observation and documentation of administrative tasks carried out by staff in each of the above job roles, followed by coding and classification; semi-structured interviews with 10 GPs, and six staff focus groups. The online survey of machine learning, artificial intelligence, and robotics experts was analysed using an ordinal Gaussian process prediction model to estimate the automatability of the observed tasks.

Results: The model predicted that roughly 44% of administrative tasks carried out by staff in general practice are “mostly” or “completely” automatable using currently available technology. Discussions with practice staff underlined the need for a cautious approach to implementation.

Conclusions: There is considerable potential to extend the use of automation in primary care, but this will require careful implementation and ongoing evaluation.

Strengths and limitations

Strengths

- This is currently the first use of a scalable measure of task-based automation applied to the work of healthcare grounded in empirical data from NHS England primary care.
- Using and Independent Bayesian Classifier Combination to achieve a single rating from multiple experts is advantageous both because it is fully Bayesian and reflects a higher

chance of accurately recovering the true automatability label of a task in an environment of uncertainty and subjectivity.

Limitations

- This study was only able to analyse administrative tasks, our use of the O*NET database in large part restricted us to looking at these administrative functions.
- Primary care looks different across all parts of the country and unfortunately, we were unable to gather data on some of these nuances and cannot be certain that we captured all the diversity.
- The subjectiveness in the task-simplification may have translated into our dataset providing very few entirely “not automatable” tasks caused, in part, by the creativity of the responder’s who would believe part of a re-arranged task to be automatable.

Introduction

Primary care in England faces numerous challenges, including increased demand, workload pressures, recruitment difficulties, and budgetary constraints [1–4]. All staff in general practice, including General Practitioners (GPs), nurses, pharmacists, and non-clinical staff, carry out many administrative tasks. There has been much discussion of the bureaucratic burden on general practice and how to streamline these tasks to free up more time for direct patient care. A report in 2015 identified several potential areas of concern, including time spent on reclaiming financial reimbursement for services rendered, entering data into various unintegrated practice-based information systems, processing information from hospitals and other external providers, keeping on top of changing requirements of commissioners, and helping patients navigate a fragmented and poorly coordinated system [5]. The authors concluded that over 50% of practice time was spent on bureaucracy and much of this was potentially avoidable. A majority of GPs nowadays describe their workload as excessive and detrimental to the delivery of high-quality care [6–8]. This phenomenon of ‘paperwork’ growth is consistent with the general upward trend in bureaucratic work in public sector occupations [9]. In general practice in England, administrative, non-clinical staff now outnumber clinical staff by a factor of 2.5 to 1 [10]. The Royal College of General Practitioners has recently called for significant investment in digital technology to transform care, including automation of routine administrative tasks [11].

Automation of administrative tasks may reduce the bureaucratic burden in primary care, but the nature and extent of its likely impact is currently unknown. In other sectors automation is typically perceived as a threat to workers, conjuring visions of mass unemployment. Previous research by

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

members of our team has classified 47% of occupations as being “high risk” for computer automation within the next couple of decades [12]. That study based its estimates of the probability of automatability on the US Department of Labor’s Occupational Information Network (O*NET) database of occupational characteristics and worker requirements [13]. Later studies applied the same probability estimates to employment data from other countries, on the assumption that the probabilities are consistent globally. The proportion of jobs estimated as ‘high risk’ ranged from 45% to more than 60% across different European countries [14–16]. However, where there are workforce shortages as in general practice, automation presents not a threat but an opportunity, potentially increasing productivity and leading to more satisfying job roles. The assumption has been that healthcare occupations are less likely to be affected because they rely heavily on interpersonal interactions, but they also include tasks that require information acquisition, information analysis, and decision making, all characteristics amenable to automation[17]. In primary care, automatable tasks include checking written documents for errors, reviewing and writing letters, creating referral letters, organising staff rotas, creating reports and maintaining records. Technologies such as natural language processing, voice recognition, text summarisation, robotic process automation, and machine learning to support the manipulation of information are all potentially applicable [18,19]. It is unlikely that implementation of automation technologies in general practice will cause large-scale unemployment of healthcare staff, but it could enhance their capacity to keep on top of administrative work while freeing up more time for patient care.

Our aim in carrying out the present study was to quantify the likely impact of automation in general practice to inform the assessment of future resource needs. This was not a future-gazing project. We were interested in the potential for automation using currently available technologies, rather than those that might become available at some point in the future.

Methods

This study deployed mixed qualitative and quantitative data collection and analysis. To help us make this assessment, author MW carried out 330 hours of ethnographic observations, interviews and focus groups in six general practices, conducted a survey of experts on automation technologies and used a predictive machine learning model to apply their assessments to primary care tasks, grouped according to the O*NET classification system. We present an overview of this process below in figure 1. We used a similar quantitative framework to that developed by Frey and Osborne [12] where the authors estimated the probability of an occupation being affected by automation by classifying tasks

according to the skills required to carry them out, such as dexterity, social perceptiveness, creativity, persuasion, negotiation, and originality. However, our approach differs in two important ways. We have assumed that automation in primary care may not impact entire occupations, but specific tasks that are commonly carried out by a variety of occupational roles. Second, and most important, we have augmented our analysis with ethnographic observations in primary care to ensure that the complexity of primary care tasks are accurately captured.

Following Frey and Osborne's methods, we used the O*NET classification of occupational tasks to assess automatability. O*NET contains over 2,000 detailed work activities (DWAs) for more than 1,000 occupations across all sectors of the economy and nearly 20,000 individual occupation-specific tasks arranged in a hierarchical structure. DWAs are specific descriptions of tasks performed by different occupations. For example, some of the DWAs include: Answer telephones to direct calls or provide information.; Diagnose medical conditions.; Document client health progress.; and Interact with patients to build rapport or provide emotional support. The O*NET system provides a useful basis for an analysis of automatability, but it had not been validated for use with British healthcare-specific occupation and task data. Our first task, therefore, was to carry out observation of primary care tasks to enable linkage to the O*NET system. The focus on task data is key because it is individual tasks rather than entire occupations that are typically automated [20–24]. Recent research suggests that occupations are best analysed as evolving combinations of detailed tasks, skills, and/or environments [23,25–27]. With more granular job data available than ever before, thousands of occupations can each be broken down into hundreds of numeric components or tasks, relating to the skills, knowledge, and abilities required to perform them. Examples given in the appendix (Appendix images 1-3) show how O*NET classifies selected occupational tasks carried out by GPs.

[Figure 1: overview of research method.]

Data collection and analysis

All fieldwork involving interactions with staff was undertaken by author MW, an experienced qualitative researcher, and approved by the Health Research Authority under project number 212952. MW contacted each primary care centre using a telephone and email to explain the study, detail what is required from centres that participate in the research and answer any questions. These conversations involved the practice manager and at least one partner General Practitioner from the practice to agree on participating in the study. Practices were recruited to assure

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

geographical diversity, variance in size, representation from multi-site super practices as well as individual single site practices, and willingness to participate in the research project.

The focus of the fieldwork was to observe the work practices of each type of staff member (listed below) and to understand the typical scope of tasks they are responsible for, and how those tasks are usually completed. When fieldwork concluded there were a total of 138 unique tasks that describe work carried out by each of the following staff roles: *Administrator, Deputy Practice Manager, GP, Healthcare Assistant, Nurse Practitioner, Pharmacy Technician, Phlebotomist, Practice Manager, Practice Nurse, Practice Pharmacist, Prescription Clerk, Receptionist, Scanning Clerk, Secretary, and Medical Summariser.*

Data collected during the fieldwork consisted of field notes, photographs and documents that elucidate tasks being observed and how different tasks are carried out at each practice. During observation of each occupational type semi structured interviews were also conducted when appropriate (i.e. the participant was not working with patients at the time or otherwise performing a task that did not allow them to respond to a question). The field researcher also attended staff coffee breaks where additional semi structured interviews occurred.

A focus group was held on the final day at each site to gather staff feedback on the researcher’s observations and to check that tasks observed were representative of their occupations. All fieldwork and focus group sessions were conducted within the practice, focus groups were held in an available staff room or conference room. Data saturation was deemed to have been reached after observations at the last practice produced no new tasks. Next, a brief survey describing 15 random tasks per occupation was sent to one of the practices for further validation to check that tasks were accurately described, and that the task descriptions were accurate.

A spreadsheet was developed by author MW detailing each observed occupation and a list of the tasks they perform. These were coded and categorised according to the frequency of task, any tools or technologies used, type of software required, other occupations in the practice that also perform the task, if the task is shared with other staff members, and any other features. Important themes were elicited and refined after a continuous process of conversation and collaborative sensemaking between the authors. All observed tasks were then matched to the equivalent task in the O*NET database. This involved reading each observed task and finding its equivalent representation of that task in O*NET, for which there was always an analogous DWA. Any differences in task

representation were discussed by the research team and the most accurate and representative task was selected. Crosswalking observed tasks to DWAs enabled us to utilize O*NETS rating of skills, knowledge, and abilities required for each task and subsequence weighting of task importance per occupation. An example of the contents of O*NET using the General Practitioner occupation is provided in appendix one. This shows the task list, top skills, knowledge, and abilities as shown through the web portal interface.

The method conventionally used to identify and predict the likelihood of automation is to hypothesize the underlying dynamics that drive automation and extrapolate into the future an unspecified number of years [12,25,28–30]. This relies on subjective forecasts that may be unreliable. Instead, we decided to ask experts to base their estimates on *current* technology capabilities, rather than speculate on the future. To obtain current estimates of the automation potential of tasks, we contacted experts in the machine learning, robotics and artificial intelligence communities through academic mailing lists and lists acquired at two machine learning conferences (specifically the Conference on Neural Information Processing Systems, 2015 and 2017). Experts verified their academic and industrial experience and (optionally) their contact information.

Each survey participant was sequentially presented with five O*NET occupations and the five tasks with the largest task-importance score for that occupation. Task importance scores are rankings of their importance to the occupation that performs the task and are provided by O*NET. Participants were asked the following question: *“Do you believe that technology exists today that could automate these tasks?”*

Each task was ranked on an ordinal scale from 1.0 to 4.0: not automatable today (1.0), mostly not automatable today (human does most of it) (2.0), could be mostly automated today (human still needed) (3.0), completely automatable today. (4.0), or unsure. Respondents were also asked to rate their overall confidence in their answers.

The multiple ratings or labels for each task were combined with Independent Bayesian Classifier Combination (IBCC), a principled Bayesian approach to combine multiple classifications [31,32]. IBCC creates a posterior prediction over labels that reflects the tendency of the individual labellers to concur with ultimately chosen label values. Crucially, it requires no forecasting or prediction by the expert participants. We believe it is the best method we have for reaching a consensus of experts' opinions. IBCC task scores were averaged into 313 work activities (note: an activity is a unique set of

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

tasks in O*NET taxonomy). Therefore, activity automatability labels concentrate around whole and half values. The final values were rounded to the closest 0.5 (a half-class). A score of 4.0 indicates a “fully automatable” work activity, while 1.0 indicates a work activity that “cannot be automated” using currently available technology [23].

Finally, all data sources were combined to form the basis of the analysis, including: 1) the validated list of primary care tasks obtained from the ethnographic observations, interviews and focus groups, coded and classified allowing us to use 2) the corresponding O*NET activity skills, knowledge, and ability numerical attributes, and 3) the inferred automatability scores of these activities informed by the expert survey. A machine learning framework was employed to infer a functional mapping between the skills, knowledge, and ability characteristics of work activities and the ground truth automatability elicited from the expert survey. The anticipated output was a flexible function estimation capable of modelling complex, non-linear relationships between the features (skills, abilities, knowledge) and the target (automatability). Function mapping is routinely used to map the input (data you have) of a set into a different output (missing data) set. Just as we map skills, knowledge, and abilities to the expert survey data, function mapping is also used to predict someone’s weight based on their height, or by using income to assess loan payments.

The collection of work activities (N=313) labelled in the expert survey of automatability was used as a training data set for the Gaussian Process Ordinal Regression model. Once trained, the model can estimate the automatability of “unlabelled” work activities, i.e. our observed healthcare tasks. This is preferable to asking experts to provide judgements on the full range of tasks observed, which would be time-consuming and prohibitively expensive. We used a Gaussian process model for two reasons. First, supervised learning problems in machine learning can be cast directly into the Gaussian process framework. Compared to traditional machine learning approaches, Gaussian processes are particularly well suited when large amounts of data – often referred to as big data – are not available or even possible, which is often required to prevent over-fitting and to ensure a good generalizability of models to new data. Second, being a non-parametric model, Gaussian processes automatically control model complexity and provide a predictive distribution over the target function, providing formal bounds on uncertainty for classification and regression tasks. A Gaussian process can be defined as a distribution over functions, mapping from our input features X of work activities to the probability of automatability.

A Kernel Density Estimation (KDE) analysis was employed to show the distribution of the inferred automatability of the healthcare tasks. KDE allows visualization of a smooth curve that represents the “shape” of the data distribution as a replacement for a discrete histogram. Further technical details of the quantitative framework and analysis, as well as a detailed description of the observational methods, are documented in a previous publication [33].

Patient and Public Involvement

No patients were involved in this study.

Results

Administrative tasks

Of the 19 practices contacted: six agreed to participate, six expressed interest and intended to participate but never followed up, and seven declined to participate due to lack of time or financial burden. Ethnographic observations were carried out in six primary care centres in England, located in Oxfordshire, Yorkshire, Berkshire, Surrey, and the West Midlands. The practice list sizes ranged from 5,000 to 24,000 patients. The researcher spent one week in each location, observing work carried out by GPs, other primary care clinicians and non-clinical staff. 330 hours of observations resulted in descriptions of 65 unique sessions of work carried out by primary care staff, each ranging from 45 minutes to 90 minutes. These observations can be best illustrated by focusing on two large and significant types of work that account for many of the administrative tasks carried out – letter work and clinical documentation. It is important to note that although we classified documentation as administrative work, we appreciate the medical importance of the tasks and that they involve several sub-tasks, some of which require specialist medical knowledge.

Letters

This type of workload is the bundle of tasks and workflows that relate to the production of letters. Letter production flows in two directions; incoming and outgoing. Incoming letters contain multiple forms of implicit and explicit information. Processing this information results in tasks such as responding to requests, storing physical and digital copies of the letter, applying a medical coding scheme, and using information contained in the letter to inform or modify patient treatment and/or other clinical decisions. Outgoing letters involved multiple processes as well, such as audio dictation, transcription of audio text, application of preformatted templates, appending directions and notes to audio files or digital documents, referencing information in the electronic medical record, archiving letters, and posting the letter. The specific details of how this work is carried out varies from practice to practice. The workflow can look remarkably different in a single-site independent

practice when compared to the workflow in a super-practice, sharing services across multiple sites. A single-site practice may allocate this work to receptionists who process letters, while a super-practice might outsource all medical coding, audio transcription, and typing to a central location.

The following workflow was observed at a small single site practice. Receptionists sorted incoming letters into groups, giving them red, amber, or green status dependent on response urgency. Then, they scanned the letters into the system and used a commercial software system called DocMan to sort them using the colour-coded system. One of the receptionists is assigned the task of “dealing” with the letters. This involves opening a letter that has been briefly coded by a GP with the highlighter tool in the document management software, reviewing the highlighted medical terms or phrases, and allocating appropriate Read codes. The letter is then archived on the system, the task is marked complete, and the receptionist continues to the next letter. If the receptionist has a question about the appropriate Read code, text comments are added to the document, which is then saved, and assigned as a task on the system, i.e. sent back to the GP. Other types of letters, for example, from a medical supplier requesting information about a patient who requires a specific device, must find the required information in the relevant health record, then use a template letter to respond to the supplier.

The length of letters ranges from precise and short to long and verbose. Letters have a different utility dependent on who is writing them, who is reading them, and the context. For example, during our research a GP was observed reading a 9-page letter from secondary care. The GP skim-read the letter on her monitor and after a few minutes declared, “ah hah, you see here it is”, highlighting a paragraph with the mouse cursor, saying “this is all I need, out of this whole document, this is what is relevant to me.” She continues: “I can tell the rest of this letter is for their [the specialist’s office] purposes, it’s a form of documentation, I don’t need all this.” The GP did not require most of the content of the letter, only a single paragraph. Her prior experience of the specialist who wrote the letter enabled her to extract the relevant information efficiently and discount the rest.

Documentation of clinical consultations

A scenario in which a system could automate clinical documentation using natural language processing and text summarisation was described to staff in focus groups and interviews, provoking lengthy discussion. Some participants saw it as a potentially useful technology, with multiple “when can we have that” comments, while others felt it might have negative consequences on the clinical encounter and could even undermine the skills and critical thinking of the clinician. One GP argued that having a transcript of the entire conversation would be a waste of time, producing a great deal

of unnecessary information. It is not uncommon for GPs to re-write or edit their notes, particularly notes for new cases or complex circumstances, or when the notes are written by trainees. What may look like a relatively simple administrative task in fact involves a considerable amount of clinical skill and experience. Removing this task from the clinical workflow would remove an opportunity for the clinician to reflect critically about an interaction. Research shows that writing engages the brain, allows GPs to be better observers, supports empathy, and engages critical thinking [34].

Scope for automation

The online survey of experts in automation systems achieved 156 responses from academic and industrial experts around the world, providing a total of 4,599 task level ratings. The distribution of respondents by country and qualification is shown in figures 2 and 3.

[Figure 2: Number of survey respondents by country.]

[Figure 3: Number of survey respondents by qualification.]

The Gaussian Process machine learning model that we applied to the data suggests that about 44% of the administrative tasks performed in primary care are “mostly” or “completely” automatable using current technology. Figure 4 shows the distribution of automatability for all administrative tasks performed by general practice staff.

[Figure 4: Distribution of Automatability Scores. Plot shows the Kernel Density Estimate of the Scores.]

As discussed above, tasks performed by various occupations overlap, but the automatability of those tasks are independent of the occupation, i.e. an automatability score is not dependent on who performs the task. For example, medical coding has an estimated automatability of 3.2. Thus, automation of this task, which is mostly performed by receptionists, summarisers, and secretaries, but also on occasion by GPs depending on the practice, could either remove this task from their workloads altogether, if the systems are considered sufficiently reliable, or reduce it to checking codes produced by an automated system. Some of the most highly automatable tasks in our data include: *payroll and managing finances, checking and sorting post, printing letters, communicating*

with patients through texting, management of paper archives (onsite or offsite), transcription, email account management, letter scanning, checking for errors in paperwork, and internal communications (for example messages to staff or new employee inductions).

To further understand the automatability predictions, three categories were generated: automatable (predicted score is >3.0), not-automatable (predicted score < 2.0), and partly-automatable (predicted score ≥ 2.0 to ≤ 3.0). Figure 5 presents the predicted automatability scores of administrative occupation tasks classified into these risk categories.

[Figure 5: Percentage of Administrative Tasks performed in Primary Care by Automatability Categories.]

Scanning clerks are typically, but not always, part time employees that are responsible for scanning physical documents (mostly letters) and creating digital objects (usually a PDF file) and attach these files to the patients' health record or import files into other document management systems. Scanning clerks also add metadata to the PDF so the files are searchable and contain all contextual information. Results show that only the work of scanning clerks could be fully automated using currently available technology, thus eliminating the need for that occupation entirely. Many routine administrative tasks and typical office work (e.g. answering phones, writing letters and managing email) are undertaken by all occupational groups, but the impact of automation will vary according to the number of tasks in each automatable category that they perform. Figures 6 and 7 below show the potential impact of automation technologies by occupation for clinical and non-clinical staff.

[Figure 6: (KDE) Distribution of Task Automatability for Primary Care Occupations (non-clinical staff).]

[Figure 7: (KDE) Distribution of Task Automatability for Primary Care Occupations (clinical staff).]

Further detailed results can be found in the technical report for The Health Foundation, currently in preparation [35].

For peer review only

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Discussion

We believe this study is the first to combine detailed validated observations of workflows and a specially developed ordinal scale of automatability to assess the potential for automating various administrative tasks routinely carried out by primary care staff. Our findings indicate considerable scope for extending the use of automation in general practice. It is important to acknowledge that certain capabilities of automation already exist in primary care. In some practices scanning clerks put paper documents into scanners that turn physical text into digital type using optical character recognition. Another form of automation is the use of templates and pre-generated text in desktop word processors or in electronic medical record notes. Automated report generation and custom search queries to input patient data to different systems is another form of automation that is sometimes used. Our analysis has looked at the scope for further automation of tasks and we conclude that it is considerable.

A few caveats must be borne in mind. Observations took place in only six general practices. These varied in terms of list size and geographical location, but we cannot claim that these were representative of all practices in England. The survey of automation experts was, we believe, the largest of its kind to date and it provides a robust baseline estimate of the extent to which tasks are automatable using only currently available technology, but participants did not have specific expertise in the application of automated systems to healthcare. We made up for this deficiency by inviting discussion and feedback from primary care staff involved in interviews and focus groups, but they, in turn, lacked direct experience of using the automated systems that might eventually modify their working practices.

Automation technologies will be most useful when they provide clinicians with the right information at the right time; employing filters of context and urgency to remove details that are irrelevant or distracting. Automation of some mundane tasks, such as telephone answering, letter writing, document scanning, email monitoring, and information filtering, could free up valuable time for patient-facing tasks. . We find that many tasks involving data entry, sending information, formatting information, maintaining records, transcription, and operating office equipment to name a few, have high potential for automation. Many practices have basic forms of automated telephone answering, but more sophisticated systems are now available that would enable a voice assistant to help triage the hundreds of phone calls that a practice can receive in a day. However, there are drawbacks to automated phone systems that would need careful consideration and evaluation. Receptionists help prevent unneeded appointments by screening phone calls and getting patients to the right area for their appointment, these skills rely on social perception and an understanding of the patient's

history. Automated phone systems must be aware of this function and help guide patients through their case without scheduling unneeded appointments.

Clinical documentation could benefit from automation for three main reasons. First, these are tasks that every clinician must complete with a degree of urgency. It is considered good medical practice to document the patient consultation as soon as possible, i.e. it cannot be left to the end of the day. Second, documentation requires attention to detail and is time intensive. Even very brief encounters require accurate documentation. Third, this work is typically squeezed into a busy schedule, usually performed during the time it takes a patient to leave the examination room and the next patient arriving. Vendors of electronic health record systems have advocated the use of automation to facilitate record-keeping by use of structured data formats [36–38]. Natural language processing can be used to dissect patient-doctor conversations and create notes [39]. How to ensure the validity and reliability of such systems has been the subject of much research and remains a controversial topic [18,40]. However, a recent survey of UK GPs revealed considerable optimism about the likelihood that clinical documentation would be fully automated in the near future, with a large majority believing this would occur within ten years [41].

Documentation of clinical consultations is a good example of why automation should be approached with caution. Automating this type of work could have unintended consequences. Documentation of clinical encounters serves as a “tool for thinking” for doctors and as an important means of coordination and communication [40]. Reading and writing clinical notes engages the medical professionals’ critical faculties in a useful way. Allowing machines to take over would circumvent the way clinicians have been trained to document clinical encounters [42]. It would be important to identify the skills medical professionals would need in order to successfully work with automation technologies while maintaining a high level of care quality.

Automation technologies can be expensive and time-consuming to implement, so those contemplating the purchase of such systems should consider the following issues: Which types of tasks would be most beneficial to the practice if automated? Are those tasks amenable to automation? If this results in a change in occupational roles, who will it affect most? How will staff be trained and supported to use the new technology? If tasks are removed from the workload of a staff member, what kind of work (if any) will replace the task? And who will be responsible for the maintenance of the tasks while new workflows are in development?

Some tasks with high automatability scores will be challenging to implement for social and organisational reasons. Previous research has shown that staff contribute to patient quality and safety through hidden ways that emerge in the moment and require tacit knowledge [43]. The tacit

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

and contextual work, such as reading social cues, currently lies outside the capabilities of automated systems and machine learning technologies [29]. Healthcare is crucially dependent on these human skills and patient care would suffer if automation is pursued to their detriment. Our observations revealed a type of work that is often driven by exceptions, interruptions, competing critical requests and time sensitivity. Work practices that may appear routine, inefficient or superfluous, often have an intrinsic value for cooperative teamwork [44]. While our quantitative findings reveal considerable potential for automation, our in-depth fieldwork demonstrated the need for a cautious approach given the complexity of social systems in healthcare delivery environments.

Observing and gathering data on tasks in primary care presents at least three major challenges: task variance, task fit, and clinical or administrative task categorization.

First, like primary care itself, tasks are greatly varied. Variance in tasks can occur in the order in which parts of the task are performed, how long they take, the occupational role of the person performing the task, the “importance” of the task or how time-critical it is, and how many individuals become involved in completing it.

There are further complexities that emerge due to the task being performed in a healthcare context, when attention is given to patients' needs, backgrounds, and particular problems. In this context many administrative tasks require specialist medical knowledge. This is what makes work in healthcare different and exceptional, when compared to other fields with similar task descriptions.

The second challenge is that the same occupation does not always perform the same set of tasks across practices. We found that some tasks that receptionists routinely perform in one practice were performed by administrators at another practice. The practices we observed differed considerably in their organisational forms, from single-site practices working independently to larger groups or “super-practices”, who share services across multiple sites. The issue of “task fit”, i.e. allocating tasks to the appropriate person, involves both matching the task to the most appropriate pay grade and making sure clinicians are shielded from unnecessary administrative work. The driving force behind the assigning of tasks to occupations in this study was to accurately represent the core identity and “scope of work” of each observed occupation.

The third issue is the fundamental difference between “clinical” and “administrative” tasks. There is an abundance of administrative and clerical tasks that occur in primary care. While the field researcher observed, catalogued, and categorized nearly every clinical interaction that takes place in the domain of primary care, it was often hard to distinguish administrative tasks from those that are

purely clinical. Also, the O*NET database that was used for our analysis provided more terminology and language to describe administrative tasks and office work than it did for nuanced clinical work.

Conclusion

Evidence from our study show there is great potential to impact workloads in primary care by automating certain administrative tasks (approximately 44%, broadly). We believe that careful introduction of currently available technologies has considerable potential to reform administrative workflows and increase productivity in primary care. However, this should not be seen as a magic bullet that will solve the workload crisis in the immediate future. Introduction of new forms of automation should proceed cautiously with ongoing evaluation of their wider impacts.

Competing interests

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: all authors had financial support from The Health Foundation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Author Contributions

Matthew Willis: Writing, figure 1, coordination of authors, literature search, citation management, study design, formatting, draft iteration, data collection, data interpretation, manuscript submission, revisions.

Paul Duckworth: Figures, draft writing, editing, data collection, data analysis, literature search, data interpretation, revisions.

Angela Coulter: Reviewing and editing, writing, data interpretation, study design, revisions.

Eric T. Meyer: Reviewing and editing, writing, data interpretation, study design, revisions.

Michael Osborne: Data analysis, data interpretation, technical advising, figures, data collection, study design.

Funding

This work was supported by The Health Foundation, Award #7559. The study funders played no role in the study design, writing of the report, or the decision to submit the manuscript for publication. Researchers were independent of influence from study funders.

Data Sharing

All data relevant to the study are included in the article.

References

1 Hopson C. The State of the NHS provider sector. 2016. <https://www.nhsproviders.org/news-blogs/news/nhs-trust-leaders-warn-staff-shortages-now-outweigh-fears-over-funding>

2 Baird B, Charles A, Honeyman M, *et al*. Understanding pressures in general practice. London, UK: 2016. <https://www.kingsfund.org.uk/publications/pressures-in-general-practice>

3 Roland M, Everington S. Tackling the crisis in general practice. *BMJ* 2016;**352**:i942. doi:10.1136/bmj.i942

4 General Practice Forward View. NHS Engl. 2016.<https://www.england.nhs.uk/gp/gpfv/about/>

5 Clay H, Stern R. Making Time in General Practice. Prim. Care Found. Alliance. 2015;;1–83.

6 Byrne L, Bottomley J, Turk A. British Medical Association Survey of GPs in England. London: 2016.

7 Gibson J, Checkland K, Coleman A, *et al*. Eighth National GP Worklife Survey. Manchester: 2016. <http://research.bmh.manchester.ac.uk/healthconomics/research/Reports/EighthNationalGPWorklifeSurveyreport/>

8 Hobbs FDR, Bankhead C, Mukhtar T, *et al*. Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007–14. *Lancet (London, England)* 2016;**387**:2323–30. doi:10.1016/S0140-6736(16)00620-6

9 Hitchcock ALK, Sundorff E. Work in Progress: Towards a leaner, smarter public-sector workforce. 2017.

10 NHS statistics, facts and figures. London, UK: 2017. <https://www.nhsconfed.org/resources/key-statistics-on-the-nhs>

11 Fit for the Future: a Vision for General Practice. London, UK: 2019. <https://www.rcgp.org.uk/policy/future-vision.aspx>

12 Frey CB, Osborne MA. The future of employment: How susceptible are jobs to computerisation? *Technol Forecast Soc Change* 2017;**114**:254–80. doi:10.1016/j.techfore.2016.08.019

13 National Center for O*NET Development. 2016.https://www.doleta.gov/programs/onet/eta_default.cfm

14 Jeremy Bowles. The computerisation of european jobs. Bruegel. 2014.<http://bruegel.org/2014/07/the-computerisation-of-european-jobs/>

15 Pajarinen M, Rouvinen P. Computerization Threatens One Third of Finnish Employment. Published Online First: 13 January 2014.<https://econpapers.repec.org/paper/rifbriefs/22.htm> (accessed 27 May 2019).

16 Knowles-Cutler A, Frey C, Osborne M. Agile Town: The Relentless March of Technology and London’s Response. London, UK: 2014. <https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/uk-futures/london-futures-agiletown.pdf>

17 Jamieson GA, Skraaning G. Levels of Automation in Human Factors Models for Automation Design: Why We Might Consider Throwing the Baby Out With the Bathwater. *J Cogn Eng Decis Mak* 2018;**12**:42–9. doi:10.1177/1555343417732856

- 18 Kreimeyer K, Foster M, Pandey A, *et al.* Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review. *J Biomed Inform* 2017;**73**:14–29. doi:10.1016/j.jbi.2017.07.012
- 19 Ng A. What Artificial Intelligence Can and Cant Do Right Now. *Harv Bus Rev* <https://hbr.org/2016/11/%0Awhat-artificial-intelligence-can-and-cant-do-right-now>
- 20 Bresnahan T, Brynjolfsson E, Hitt L. Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-Level Evidence. Cambridge, MA: 1999. doi:10.3386/w7136
- 21 Brynjolfsson E, McAfee A. *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy.* 2012.
- 22 Brynjolfsson E, McAfee A. *The second machine age : work, progress, and prosperity in a time of brilliant technologies.* W. W. Norton & Company 2014.
- 23 Duckworth P, Graham L, Osborne M. Inferring Work Task Automatability from AI Expert Evidence. In: *ACM conference on Artificial Intelligence, Ethics, and Society.* Honolulu, Hawaii: 2019. <http://www.aies-conference.com/accepted-papers/>
- 24 Susskind RE, Susskind D. *The future of the professions: how technology will transform the work of human experts.* OUP Oxford 2017.
- 25 Acemoglu D, Restrepo P. Robots and Jobs: Evidence From Us Labor Markets. *Nber Work Pap Ser* 2017.
- 26 Arntz M, Gregory T, Zierahn U. Revisiting the risk of automation. *Econ Lett* 2017;**159**:157–60. doi:10.1016/j.econlet.2017.07.001
- 27 James Manyika, Michael Chui, Mehdi Miremadi, *et al.* Harnessing automation for a future that works. McKinsey Glob. Inst. 2017. <http://www.mckinsey.com/global-themes/digital-disruption/harnessing-automation-for-a-future-that-works>
- 28 Autor DH, Handel MJ. Putting Tasks to the Test : Human Capital, Job Tasks, and Wages. *J Labor Econ* 2013;**31**:59–96. doi:10.1086/669332
- 29 Bakhshi H, Downing JM, Osborne MA, *et al.* The Future of Skills: Employment in 2030. 2017.
- 30 Goos M, Manning A, Salomons A. Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *Am Econ Rev* 2014;**104**:2509–26. doi:10.1257/aer.104.8.2509
- 31 Kim H-C, Ghahramani Z. Bayesian Classifier Combination. In: *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics.* PMLR 2012. 619–627. <http://proceedings.mlr.press/v22/kim12.html>
- 32 Simpson E, Roberts S, Psorakis I, *et al.* Dynamic Bayesian Combination of Multiple Imperfect Classifiers. Published Online First: 8 June 2012. <http://arxiv.org/abs/1206.1831>
- 33 Willis M, Duckworth P, Coulter A, *et al.* The Future of Health Care: Protocol for Measuring the Potential of Task Automation Grounded in the National Health Service Primary Care System. *JMIR Res Protoc* 2019;**8**:e11232. doi:10.2196/11232
- 34 Wald HS, Borkan JM, Taylor JS, *et al.* Fostering and Evaluating Reflective Capacity in Medical Education: Developing the REFLECT Rubric for Assessing Reflective Writing. *Acad Med* 2012;**87**:41–50. doi:10.1097/ACM.0b013e31823b55fa
- 35 Willis M, Duckworth P, Graham L, *et al.* Future of Healthcare: Computerisation, Automation,

and General Practice Services. London, UK: 2019.

36 Rule A, Rick S, Chiu M, *et al.* Validating free-text order entry for a note-centric EHR. *AMIA . Annu Symp proceedings AMIA Symp* 2015;**2015**:1103–10.<http://www.ncbi.nlm.nih.gov/pubmed/26958249> (accessed 9 Sep 2018).

37 Lin SY, Shanafelt TD, Asch SM. Reimagining Clinical Documentation With Artificial Intelligence. *Mayo Clin Proc* 2018;**93**:563–5. doi:10.1016/j.mayocp.2018.02.016

38 Verghese A, Shah NH, Harrington RA. What This Computer Needs Is a Physician: Humanism and Artificial Intelligence. *JAMA* 2018;**319**:19. doi:10.1001/jama.2017.19198

39 Klann JG, Szolovits P. An intelligent listening framework for capturing encounter notes from a doctor-patient dialog. *BMC Med Inform Decis Mak* 2009;**9 Suppl 1**:S3. doi:10.1186/1472-6947-9-S1-S3

40 Bansler JP, Havn EC, Schmidt K, *et al.* Cooperative Epistemic Work in Medical Practice: An Analysis of Physicians’ Clinical Notes. *Comput Support Coop Work* 2016;**25**:503–46. doi:10.1007/s10606-016-9261-x

41 Blease C, Bernstein MH, Gaab J, *et al.* Computerization and the future of primary care: A survey of general practitioners in the UK. *PLoS* 2018;**13**:1–12. doi:10.1371/journal.pone.0207418 December

42 Willis M, Jarrahi MH. Automating Documentation: A critical perspective into the role of artificial intelligence in clinical documentation. In: *iConference*. Washington, D.C.: : Springer 2019.

43 Swinglehurst D, Greenhalgh T, Russell J, *et al.* Receptionist input to quality and safety in repeat prescribing in UK general practice: ethnographic case study. *BMJ* 2011;**343**:d6788. doi:10.1136/bmj.d6788

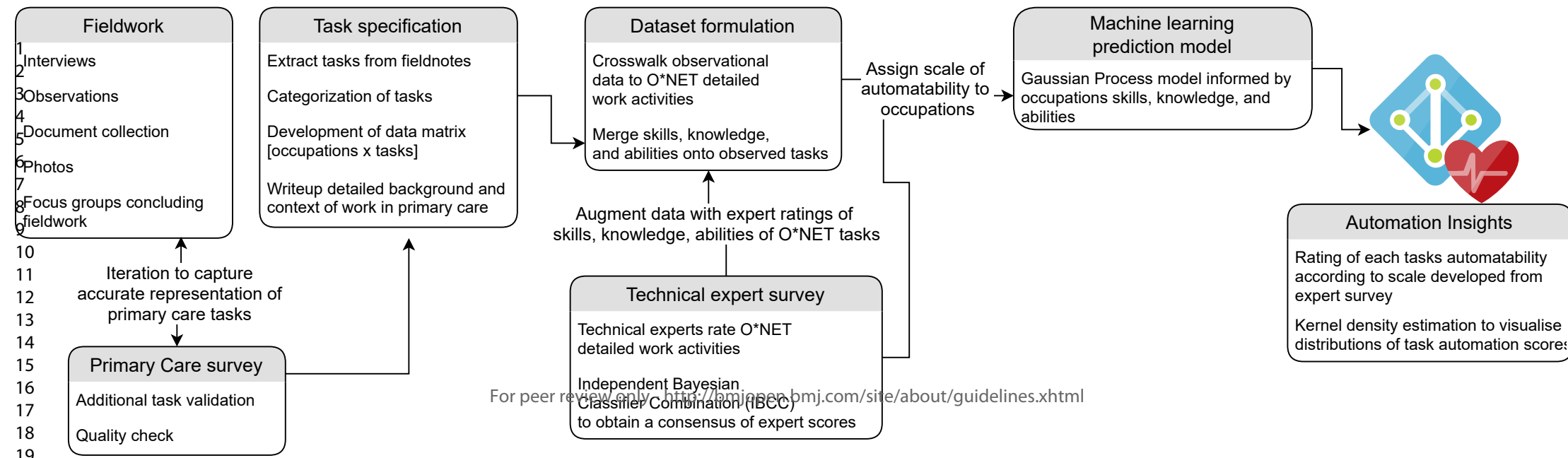
44 Cabitza F, Ellingsen G, Locoro A, *et al.* Repetita iuvant : Exploring and Supporting Redundancy in Hospital Practices. *Comput Support Coop Work* 2018.

Appendix

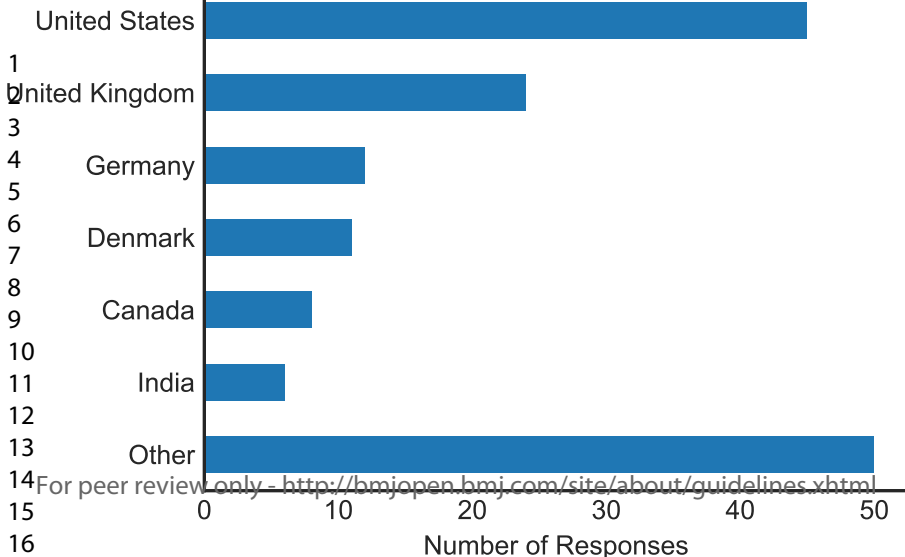
[Appendix Image 1: General Practitioner task list from O*NET database displayed in web portal report.]

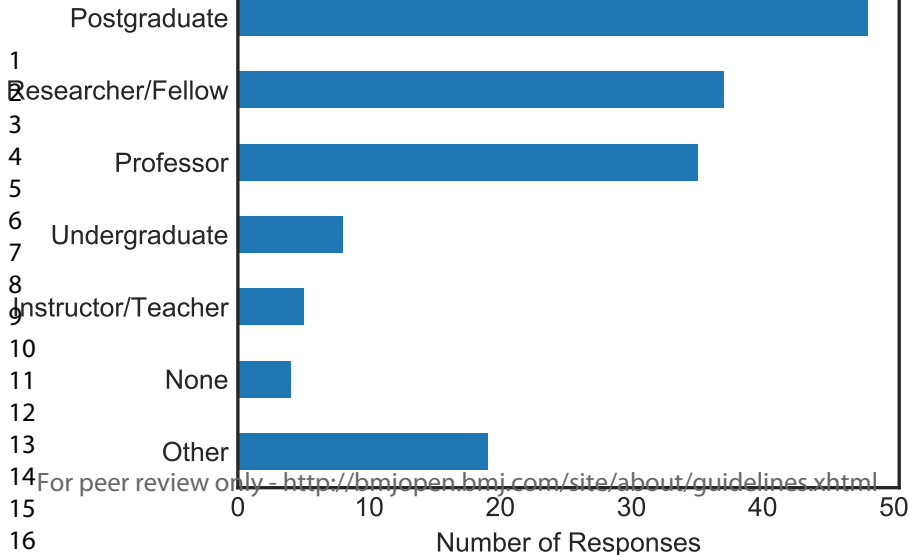
[Appendix Image 2: General Practitioner top knowledge and skills from O*NET database displayed in web portal report.]

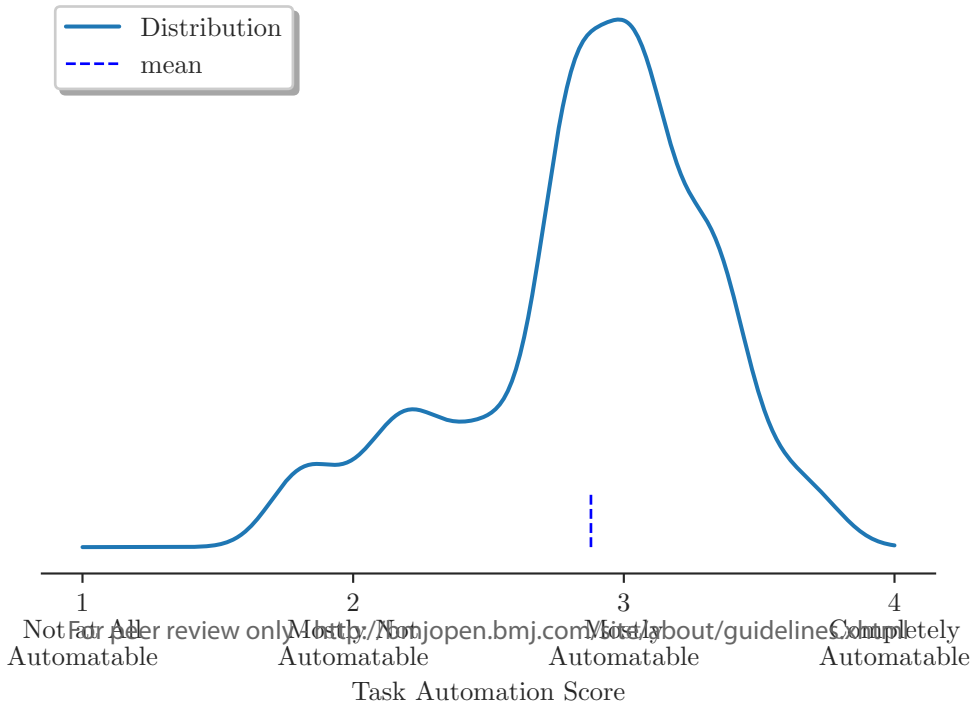
[Appendix Image 3: General Practitioner top abilities and detailed work activities from O*NET database displayed in web portal report.]



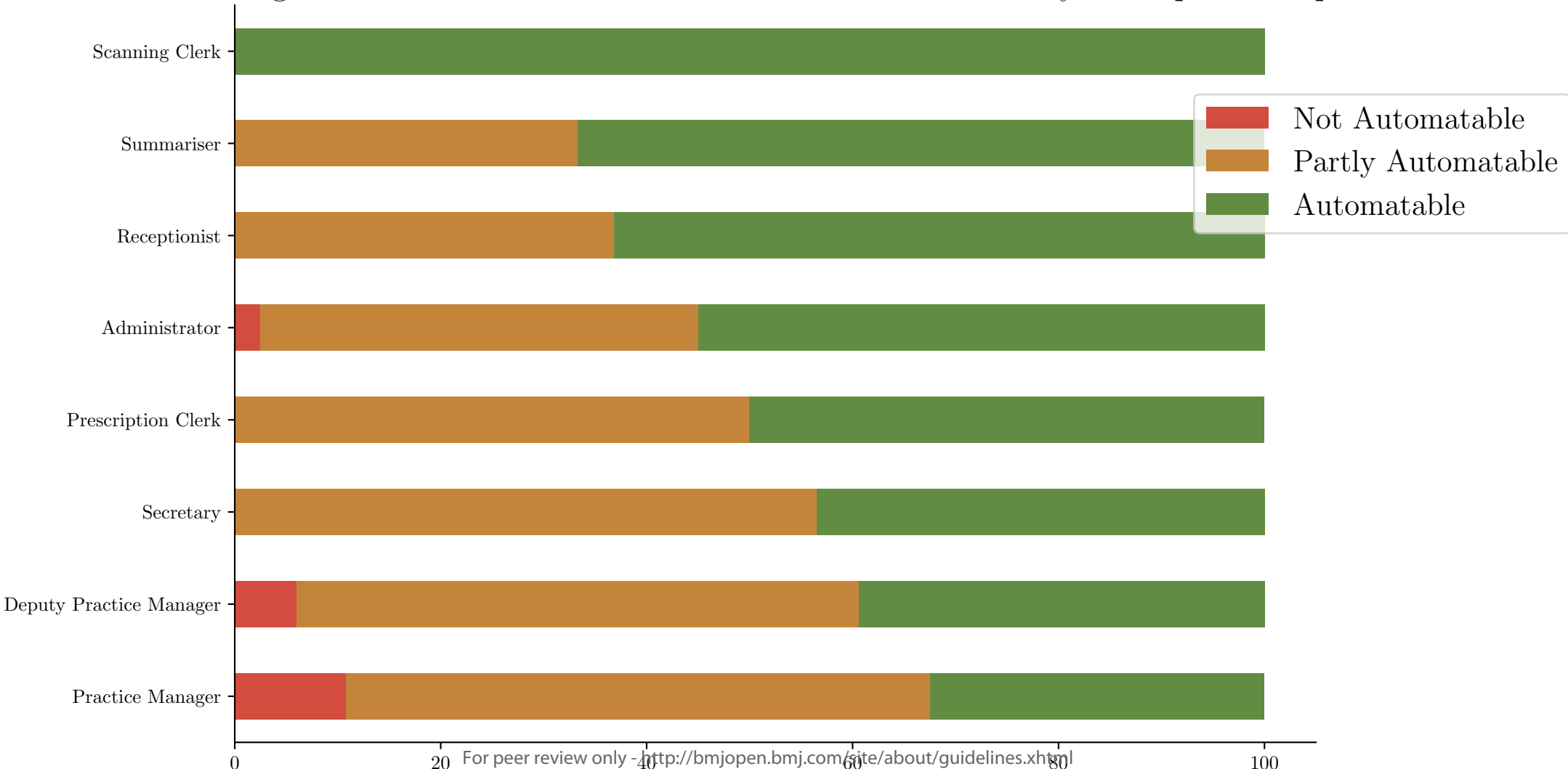
Top Responder Locations



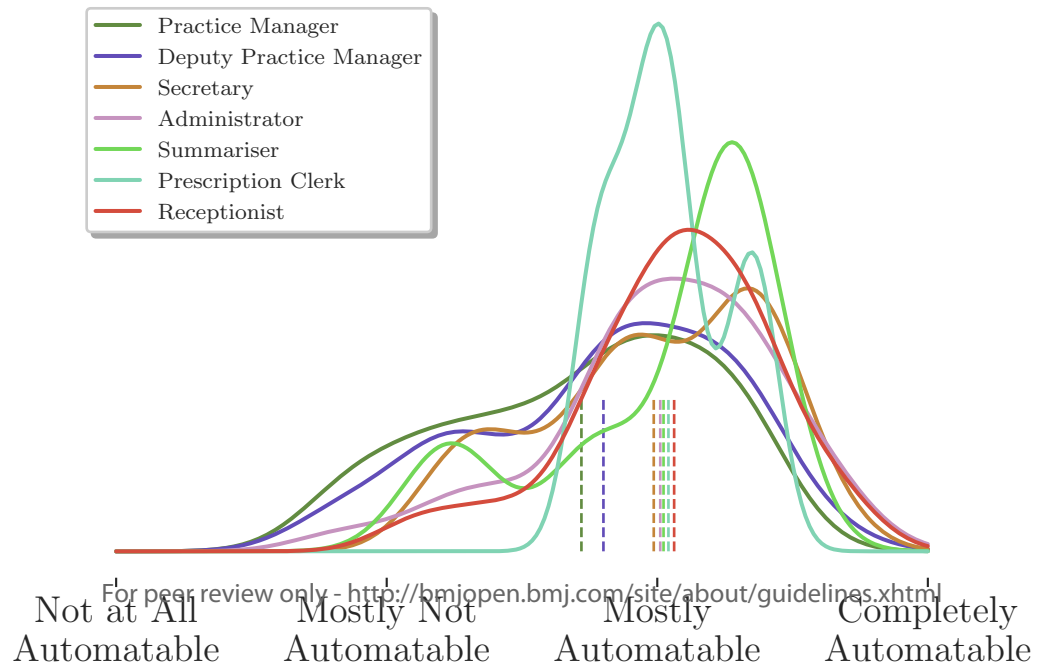




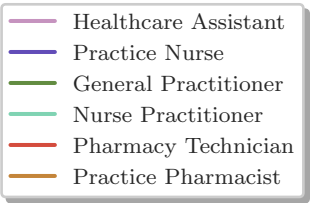
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34



Percentage of Tasks Performed by Automatable Category



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20



Not at All Automatable Mostly Not Automatable Mostly Automatable Completely Automatable

Physicians who diagnose, treat, and help prevent diseases and injuries that commonly occur in the general population. May refer patients to medical specialists when needed for further diagnosis or treatment.

Sample of reported job titles: Board Certified Family Physician, Family Medicine Physician, Family Physician, Family Practice Medical Doctor (FP MD), Family Practice Physician, Family Practitioner, Medical Doctor (MD), Medical Staff Physician, Physician, Primary Care Physician

View report: [Summary](#) [Details](#) [Custom](#)

[Knowledge](#) | [Skills](#) | [Abilities](#) | [Detailed Work Activities](#)

Save Table ([XLS](#)/[CSV](#))

All 10 displayed

Importance	Category	Task
95	Core	⊕ Prescribe or administer treatment, therapy, medication, vaccination, and other specialized medical care to treat or prevent illness, disease, or injury.
93	Core	⊕ Order, perform, and interpret tests and analyze records, reports, and examination information to diagnose patients' condition.
93	Core	⊕ Collect, record, and maintain patient information, such as medical history, reports, or examination results.
92	Core	⊕ Monitor patients' conditions and progress and reevaluate treatments as necessary.
91	Core	⊕ Explain procedures and discuss test results or prescribed treatments with patients.
88	Core	⊕ Advise patients and community members concerning diet, activity, hygiene, and disease prevention.
85	Core	⊕ Refer patients to medical specialists or other practitioners when necessary.
77	Core	⊕ Coordinate work with nurses, social workers, rehabilitation therapists, pharmacists, psychologists, and other health care providers.
75	Core	⊕ Direct and coordinate activities of nurses, students, assistants, specialists, therapists, and other medical staff.

Importance Knowledge

100		⊕ Medicine and Dentistry — Knowledge of the information and techniques needed to diagnose and treat human injuries, diseases, and deformities. This includes symptoms, treatment alternatives, drug properties and interactions, and preventive health-care measures.
91		⊕ Psychology — Knowledge of human behavior and performance; individual differences in ability, personality, and interests; learning and motivation; psychological research methods; and the assessment and treatment of behavioral and affective disorders.
86		⊕ Customer and Personal Service — Knowledge of principles and processes for providing customer and personal services. This includes customer needs assessment, meeting quality standards for services, and evaluation of customer satisfaction.
85		⊕ Biology — Knowledge of plant and animal organisms, their tissues, cells, functions, interdependencies, and interactions with each other and the environment.
83		⊕ Therapy and Counseling — Knowledge of principles, methods, and procedures for diagnosis, treatment, and rehabilitation of physical and mental dysfunctions, and for career counseling and guidance.

[back to top](#)

Skills [Save Table \(XLS/CSV\)](#)

10 of 11 displayed

Importance Skill

85		⊕ Critical Thinking — Using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions or approaches to problems.
81		⊕ Active Listening — Giving full attention to what other people are saying, taking time to understand the points being made, asking questions as appropriate, and not interrupting at inappropriate times.
81		⊕ Judgment and Decision Making — Considering the relative costs and benefits of potential actions to choose the most appropriate one.
78		⊕ Reading Comprehension — Understanding written sentences and paragraphs in work related documents.
78		⊕ Writing — Communicating effectively in writing as appropriate for the needs of the audience.
75		⊕ Complex Problem Solving — Identifying complex problems and reviewing related information to develop and evaluate options and implement solutions.

	Importance	Ability
100		⊕ Problem Sensitivity — The ability to tell when something is wrong or is likely to go wrong. It does not involve solving the problem, only recognizing there is a problem.
94		⊕ Oral Expression — The ability to communicate information and ideas in speaking so others will understand.
91		⊕ Oral Comprehension — The ability to listen to and understand information and ideas presented through spoken words and sentences.
81		⊕ Inductive Reasoning — The ability to combine pieces of information to form general rules or conclusions (includes finding a relationship among seemingly unrelated events).
81		⊕ Written Comprehension — The ability to read and understand information and ideas presented in writing.
78		⊕ Deductive Reasoning — The ability to apply general rules to specific problems to produce answers that make sense.
75		⊕ Written Expression — The ability to communicate information and ideas in writing so others will understand.
72		⊕ Speech Clarity — The ability to speak clearly so others can understand you.

[Back to top](#)

Detailed Work Activities

Save Table (XLS/CSV)

10 of 21 displayed

- ⊕ Immunize patients.
- ⊕ Prescribe medications.
- ⊕ Prescribe treatments or therapies.
- ⊕ Analyze test data or images to inform diagnosis or treatment.
- ⊕ Collect medical information from patients, family members, or other medical professionals.
- ⊕ Order medical diagnostic or clinical tests.
- ⊕ Record patient medical histories.
- ⊕ Monitor patient progress or responses to treatments.
- ⊕ Explain medical procedures or test results to patients or family members.
- ⊕ Advise communities or institutions regarding health or safety issues.

Standards for Reporting Qualitative Research (SRQR)*

<http://www.equator-network.org/reporting-guidelines/srqr/>

Authors note: this study was a mixed methods interdisciplinary project. The results reported in this article are the quantitative findings that used qualitative ethnographic fieldwork to fine tune the quantitative model and ground it in actual observations. As such, some items on this checklist may not apply. The detailed qualitative findings, coding, analysis, and theorization are a separate project and will be reported in a follow up article.

Page/line no(s).

Title and abstract

Title - Concise description of the nature and topic of the study Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended	1
Abstract - Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	2

Introduction

Problem formulation - Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	3 and 4
Purpose or research question - Purpose of the study and specific objectives or questions	4

Methods

Qualitative approach and research paradigm - Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., postpositivist, constructivist/ interpretivist) is also recommended; rationale**	4 and 5
Researcher characteristics and reflexivity - Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	5
Context - Setting/site and salient contextual factors; rationale**	9
Sampling strategy - How and why research participants, documents, or events were selected; criteria for deciding when no further sampling was necessary (e.g., sampling saturation); rationale**	5 and 6
Ethical issues pertaining to human subjects - Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	5

Data collection methods - Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and modification of procedures in response to evolving study findings; rationale**	6 and 9
Data collection instruments and technologies - Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	6 and 7
Units of study - Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	9
Data processing - Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/de-identification of excerpts	6
Data analysis - Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale**	6 and 7
Techniques to enhance trustworthiness - Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale**	6

Results/findings

Synthesis and interpretation - Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	9-11
Links to empirical data - Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	9-11

Discussion

Integration with prior work, implications, transferability, and contribution(s) to the field - Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/generalizability; identification of unique contribution(s) to scholarship in a discipline or field	14-17
Limitations - Trustworthiness and limitations of findings	2 and 16

Other

Conflicts of interest - Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	17
Funding - Sources of funding and other support; role of funders in data collection, interpretation, and reporting	17

*The authors created the SRQR by searching the literature to identify guidelines, reporting standards, and critical appraisal criteria for qualitative research; reviewing the reference lists of retrieved sources; and contacting experts to gain feedback. The SRQR aims to improve the transparency of all aspects of qualitative research by providing clear standards for reporting qualitative research.

**The rationale should briefly discuss the justification for choosing that theory, approach, method, or technique rather than other options available, the assumptions and limitations implicit in those choices, and how those choices influence study conclusions and transferability. As appropriate, the rationale for several items might be discussed together.

Reference:

O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. **Standards for reporting qualitative research: a synthesis of recommendations.** *Academic Medicine*, Vol. 89, No. 9 / Sept 2014
DOI: 10.1097/ACM.0000000000000388

BMJ Open

A qualitative and quantitative approach to assess of the potential for automating administrative tasks in general practice

Journal:	<i>BMJ Open</i>
Manuscript ID	bmjopen-2019-032412.R2
Article Type:	Original research
Date Submitted by the Author:	03-Mar-2020
Complete List of Authors:	Willis, Matthew; University of Oxford, Oxford Internet Institute; University of Michigan, School of Information Duckworth, Paul; University of Oxford, Department of Engineering Science Coulter, Angela; University of Oxford, Nuffield Department of Population Health Meyer, Eric; University of Oxford, Oxford Internet Institute; University of Texas at Austin, School of Information Osborne, Michael; University of Oxford, Department of Engineering Science
Primary Subject Heading:	General practice / Family practice
Secondary Subject Heading:	Health informatics
Keywords:	automation, office, machine learning, PRIMARY CARE, Health policy < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Organisation of health services < HEALTH SERVICES ADMINISTRATION & MANAGEMENT

SCHOLARONE™
Manuscripts



I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our [licence](#).

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which [Creative Commons](#) licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Title

A qualitative and quantitative approach to assess of the potential for automating administrative tasks in general practice

Authors

Matthew Willis^{1,5} (Corresponding Author Matthew.willis@oii.ox.ac.uk, +44 01865 287 213); Paul Duckworth²; Angela Coulter³; Eric T. Meyer^{1,4}; Michael Osborne²

1. University of Oxford
Oxford Internet Institute
1 st. Giles
Oxford OX1 3JS
United Kingdom
2. University of Oxford
Department of Engineering Science
Parks Road
Oxford OX1 3PJ
United Kingdom
3. University of Oxford
Nuffield Department of Population Health
Richard Doll Building
Old Road Campus
Oxford OX3 7LF
United Kingdom
4. University of Texas at Austin
School of Information
1616 Guadalupe St
Austin, TX 78701
USA
5. University of Michigan at Ann Arbor
School of Information
105 S State St
Ann Arbor, MI 48109
USA

Keywords

Automation, Office (MeSH Unique ID: D009816); Machine Learning (MeSH Unique ID: D000069550); Primary Health Care (MeSH Unique ID: D011320); Health Policy (MeSH Unique ID: D006291); Organization and Administration (MeSH Unique ID: D009934)

Word Count

5,180

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

Objective: To identify the extent to which administrative tasks carried out by primary care staff in general practice could be automated.

Design: A mixed-method design including ethnographic case studies, focus groups, interviews and an online survey of automation experts.

Setting: Three urban and three rural general practice health centres in England selected for differences in list size and organisational characteristics.

Participants: Observation and interviews with 65 primary care staff in the following job roles: administrator, manager, general practitioner, healthcare assistant, nurse practitioner, pharmacy technician, phlebotomist, practice nurse, pharmacist, prescription clerk, receptionist, scanning clerk, secretary, and medical summariser; together with a survey of 156 experts in automation technologies.

Methods: 330 hours of ethnographic observation and documentation of administrative tasks carried out by staff in each of the above job roles, followed by coding and classification; semi-structured interviews with 10 GPs, and six staff focus groups. The online survey of machine learning, artificial intelligence, and robotics experts was analysed using an ordinal Gaussian process prediction model to estimate the automatability of the observed tasks.

Results: The model predicted that roughly 44% of administrative tasks carried out by staff in general practice are “mostly” or “completely” automatable using currently available technology. Discussions with practice staff underlined the need for a cautious approach to implementation.

Conclusions: There is considerable potential to extend the use of automation in primary care, but this will require careful implementation and ongoing evaluation.

Strengths and limitations

Strengths

- This is currently the first use of a scalable measure of task-based automation applied to the work of healthcare grounded in empirical data from NHS England primary care.
- Using and Independent Bayesian Classifier Combination to achieve a single rating from multiple experts is advantageous both because it is fully Bayesian and reflects a higher

chance of accurately recovering the true automatability label of a task in an environment of uncertainty and subjectivity.

Limitations

- This study was only able to analyse administrative tasks, our use of the O*NET database in large part restricted us to looking at these administrative functions.
- Primary care looks different across all parts of the country and unfortunately, we were unable to gather data on some of these nuances and cannot be certain that we captured all the diversity.
- The subjectiveness in the task-simplification may have translated into our dataset providing very few entirely “not automatable” tasks caused, in part, by the creativity of the responder’s who would believe part of a re-arranged task to be automatable.

Introduction

Primary care in England faces numerous challenges, including increased demand, workload pressures, recruitment difficulties, and budgetary constraints [1–4]. All staff in general practice, including General Practitioners (GPs), nurses, pharmacists, and non-clinical staff, carry out many administrative tasks. There has been much discussion of the bureaucratic burden on general practice and how to streamline these tasks to free up more time for direct patient care. A report in 2015 identified several potential areas of concern, including time spent on reclaiming financial reimbursement for services rendered, entering data into various unintegrated practice-based information systems, processing information from hospitals and other external providers, keeping on top of changing requirements of commissioners, and helping patients navigate a fragmented and poorly coordinated system [5]. The authors concluded that over 50% of practice time was spent on bureaucracy and much of this was potentially avoidable. A majority of GPs nowadays describe their workload as excessive and detrimental to the delivery of high-quality care [6–8]. This phenomenon of ‘paperwork’ growth is consistent with the general upward trend in bureaucratic work in public sector occupations [9]. In general practice in England, administrative, non-clinical staff now outnumber clinical staff by a factor of 2.5 to 1 [10]. The Royal College of General Practitioners has recently called for significant investment in digital technology to transform care, including automation of routine administrative tasks [11].

Automation of administrative tasks may reduce the bureaucratic burden in primary care, but the nature and extent of its likely impact is currently unknown. In other sectors automation is typically perceived as a threat to workers, conjuring visions of mass unemployment. Previous research by

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

members of our team has classified 47% of occupations as being “high risk” for computer automation within the next couple of decades [12]. That study based its estimates of the probability of automatability on the US Department of Labor’s Occupational Information Network (O*NET) database of occupational characteristics and worker requirements [13]. O*NET is a large database that has been in development since 1997 and contains extensive information on a given job and the attributes of the worker, such as: personal requirements (skills and knowledge), personal characteristics (abilities and values), experience requirements (training, licensing, and experience), requirements of the job (physical, social, and organizational context of the work), occupation-specific information (title, tasks, technology skills, and tools) and the outlook of the occupation in the labor market. O*NET enables the understanding of what is expected from a worker and the detailed tasks for which the worker is responsible. Later studies applied the same probability estimates to employment data from other countries, on the assumption that the probabilities are consistent globally. The proportion of jobs estimated as ‘high risk’ ranged from 45% to more than 60% across different European countries [14–16]. However, where there are workforce shortages as in general practice, automation presents not a threat but an opportunity, potentially increasing productivity and leading to more satisfying job roles. The assumption has been that healthcare occupations are less likely to be affected because they rely heavily on interpersonal interactions, but they also include tasks that require information acquisition, information analysis, and decision making, all characteristics amenable to automation[17]. In primary care, automatable tasks include checking written documents for errors, reviewing and writing letters, creating referral letters, organising staff rotas, creating reports and maintaining records. Technologies such as natural language processing, voice recognition, text summarisation, robotic process automation, and machine learning to support the manipulation of information are all potentially applicable [18,19]. It is unlikely that implementation of automation technologies in general practice will cause large-scale unemployment of healthcare staff, but it could enhance their capacity to keep on top of administrative work while freeing up more time for patient care.

Our aim in carrying out the present study was to quantify the likely impact of automation in general practice to inform the assessment of future resource needs. This was not a future-gazing project. We were interested in the potential for automation using currently available technologies, rather than those that might become available at some point in the future.

Methods

This study deployed mixed qualitative and quantitative data collection and analysis. To help us make this assessment, author MW carried out 330 hours of ethnographic observations, interviews and focus groups in six general practices, conducted a survey of experts on automation technologies and used a predictive machine learning model to apply their assessments to primary care tasks, grouped according to the O*NET classification system. We present an overview of this process below in figure 1. We used a similar quantitative framework to that developed by Frey and Osborne [12] where the authors estimated the probability of an occupation being affected by automation by classifying tasks according to the skills required to carry them out, such as dexterity, social perceptiveness, creativity, persuasion, negotiation, and originality. However, our approach differs in two important ways. We have assumed that automation in primary care may not impact entire occupations, but specific tasks that are commonly carried out by a variety of occupational roles. Second, and most important, we have augmented our analysis with ethnographic observations in primary care to ensure that the complexity of primary care tasks is accurately captured.

Following Frey and Osborne's methods, we used the O*NET classification of occupational tasks to assess automatability. O*NET contains over 2,000 detailed work activities (DWAs) for more than 1,000 occupations across all sectors of the economy and nearly 20,000 individual occupation-specific tasks arranged in a hierarchical structure. DWAs are specific descriptions of tasks performed by different occupations. For example, some of the DWAs include: Answer telephones to direct calls or provide information.; Diagnose medical conditions.; Document client health progress.; and Interact with patients to build rapport or provide emotional support. The O*NET system provides a useful basis for an analysis of automatability, but it had not been validated for use with British healthcare-specific occupation and task data. Our first task, therefore, was to carry out observation of primary care tasks to enable linkage to the O*NET system. The focus on task data is key because it is individual tasks rather than entire occupations that are typically automated [20–24]. Recent research suggests that occupations are best analysed as evolving combinations of detailed tasks, skills, and/or environments [23,25–27]. With more granular job data available than ever before, thousands of occupations can each be broken down into hundreds of numeric components or tasks, relating to the skills, knowledge, and abilities required to perform them. Examples given in the appendix (Appendix images 1-3) show how O*NET classifies selected occupational tasks carried out by GPs.

[Figure 1: overview of research method.]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Data collection and analysis

This study received a favourable opinion from the North East Tyne & Wear South Research Ethics Committee and approval from the Health Research Authority under project number 212952. All fieldwork involving interactions with staff was undertaken by author MW, an experienced qualitative researcher. MW contacted each primary care centre using a telephone and email to explain the study, detail what is required from centres that participate in the research and answer any questions. These conversations involved the practice manager and at least one partner General Practitioner from the practice to agree on participating in the study. Practices were recruited to assure geographical diversity, variance in size, representation from multi-site super practices as well as individual single site practices, and willingness to participate in the research project.

The focus of the fieldwork was to observe the work practices of each type of staff member (listed below) and to understand the typical scope of tasks they are responsible for, and how those tasks are usually completed. When fieldwork concluded there were a total of 138 unique tasks that describe work carried out by each of the following staff roles: *Administrator, Deputy Practice Manager, GP, Healthcare Assistant, Nurse Practitioner, Pharmacy Technician, Phlebotomist, Practice Manager, Practice Nurse, Practice Pharmacist, Prescription Clerk, Receptionist, Scanning Clerk, Secretary, and Medical Summariser.*

Data collected during the fieldwork consisted of field notes, photographs and documents that elucidate tasks being observed and how different tasks are carried out at each practice. During observation of each occupational type semi structured interviews were also conducted when appropriate (i.e. the participant was not working with patients at the time or otherwise performing a task that did not allow them to respond to a question). The field researcher also attended staff coffee breaks where additional semi structured interviews occurred.

A focus group was held on the final day at each site to gather staff feedback on the researcher’s observations and to check that tasks observed were representative of their occupations. All fieldwork and focus group sessions were conducted within the practice, focus groups were held in an available staff room or conference room. Data saturation was deemed to have been reached after observations at the last practice produced no new tasks. Next, a brief survey describing 15 random tasks per occupation was sent to one of the practices for further validation to check that tasks were accurately described, and that the task descriptions were accurate.

A spreadsheet was developed by author MW detailing each observed occupation and a list of the tasks they perform. These were coded and categorised according to the frequency of task, any tools or technologies used, type of software required, other occupations in the practice that also perform the task, if the task is shared with other staff members, and any other features. Important themes were elicited and refined after a continuous process of conversation and collaborative sensemaking between the authors. All observed tasks were then matched to the equivalent task in the O*NET database. This involved reading each observed task and finding its equivalent representation of that task in O*NET, for which there was always an analogous DWA. Any differences in task representation were discussed by the research team and the most accurate and representative task was selected. Crosswalking observed tasks to DWAs enabled us to utilize O*NETS rating of skills, knowledge, and abilities required for each task and subsequence weighting of task importance per occupation. An example of the contents of O*NET using the General Practitioner occupation is provided in appendix one. This shows the task list, top skills, knowledge, and abilities as shown through the web portal interface.

The method conventionally used to identify and predict the likelihood of automation is to hypothesize the underlying dynamics that drive automation and extrapolate into the future an unspecified number of years [12,25,28–30]. This relies on subjective forecasts that may be unreliable. Instead, we decided to ask experts to base their estimates on *current* technology capabilities, rather than speculate on the future. To obtain current estimates of the automation potential of tasks, we contacted experts in the machine learning, robotics and artificial intelligence communities through academic mailing lists and lists acquired at two machine learning conferences (specifically the Conference on Neural Information Processing Systems, 2015 and 2017). Experts verified their academic and industrial experience and (optionally) their contact information.

Each survey participant was sequentially presented with five O*NET occupations and the five tasks with the largest task-importance score for that occupation. Task importance scores are rankings of their importance to the occupation that performs the task and are provided by O*NET. Participants were asked the following question: *“Do you believe that technology exists today that could automate these tasks?”*

Each task was ranked on an ordinal scale from 1.0 to 4.0: not automatable today (1.0), mostly not automatable today (human does most of it) (2.0), could be mostly automated today (human still

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

needed) (3.0), completely automatable today. (4.0), or unsure. Respondents were also asked to rate their overall confidence in their answers.

The multiple ratings or labels for each task were combined with Independent Bayesian Classifier Combination (IBCC), a principled Bayesian approach to combine multiple classifications [31,32]. IBCC creates a posterior prediction over labels that reflects the tendency of the individual labellers to concur with ultimately chosen label values. Crucially, it requires no forecasting or prediction by the expert participants. We believe it is the best method we have for reaching a consensus of experts' opinions. IBCC task scores were averaged into 313 work activities (note: an activity is a unique set of tasks in O*NET taxonomy). Therefore, activity automatability labels concentrate around whole and half values. The final values were rounded to the closest 0.5 (a half-class). A score of 4.0 indicates a "fully automatable" work activity, while 1.0 indicates a work activity that "cannot be automated" using currently available technology [23].

Finally, all data sources were combined to form the basis of the analysis, including: 1) the validated list of primary care tasks obtained from the ethnographic observations, interviews and focus groups, coded and classified allowing us to use 2) the corresponding O*NET activity skills, knowledge, and ability numerical attributes, and 3) the inferred automatability scores of these activities informed by the expert survey. A machine learning framework was employed to infer a functional mapping between the skills, knowledge, and ability characteristics of work activities and the ground truth automatability elicited from the expert survey. The anticipated output was a flexible function estimation capable of modelling complex, non-linear relationships between the features (skills, abilities, knowledge) and the target (automatability). Function mapping is routinely used to map the input (data you have) of a set into a different output (missing data) set. Just as we map skills, knowledge, and abilities to the expert survey data, function mapping is also used to predict someone's weight based on their height, or by using income to assess loan payments.

The collection of work activities (N=313) labelled in the expert survey of automatability was used as a training data set for the Gaussian Process Ordinal Regression model. Once trained, the model can estimate the automatability of "unlabelled" work activities, i.e. our observed healthcare tasks. This is preferable to asking experts to provide judgements on the full range of tasks observed, which would be time-consuming and prohibitively expensive. We used a Gaussian process model for two reasons. First, supervised learning problems in machine learning can be cast directly into the Gaussian process framework. Compared to traditional machine learning approaches, Gaussian processes are

particularly well suited when large amounts of data – often referred to as big data – are not available or even possible, which is often required to prevent over-fitting and to ensure a good generalizability of models to new data. Second, being a non-parametric model, Gaussian processes automatically control model complexity and provide a predictive distribution over the target function, providing formal bounds on uncertainty for classification and regression tasks. A Gaussian process can be defined as a distribution over functions, mapping from our input features X of work activities to the probability of automatability.

A Kernel Density Estimation (KDE) analysis was employed to show the distribution of the inferred automatability of the healthcare tasks. KDE allows visualization of a smooth curve that represents the “shape” of the data distribution as a replacement for a discrete histogram. Further technical details of the quantitative framework and analysis, as well as a detailed description of the observational methods, are documented in a previous publication [33].

Patient and Public Involvement

No patients were involved in this study.

Results

Administrative tasks

Of the 19 practices contacted: six agreed to participate, six expressed interest and intended to participate but never followed up, and seven declined to participate due to lack of time or financial burden. Ethnographic observations were carried out in six primary care centres in England, located in Oxfordshire, Yorkshire, Berkshire, Surrey, and the West Midlands. The practice list sizes ranged from 5,000 to 24,000 patients. The researcher spent one week in each location, observing work carried out by GPs, other primary care clinicians and non-clinical staff. 330 hours of observations resulted in descriptions of 65 unique sessions of work carried out by primary care staff, each ranging from 45 minutes to 90 minutes. These observations can be best illustrated by focusing on two large and significant types of work that account for many of the administrative tasks carried out – letter work and clinical documentation. It is important to note that although we classified documentation as administrative work, we appreciate the medical importance of the tasks and that they involve several sub-tasks, some of which require specialist medical knowledge.

Letters

This type of workload is the bundle of tasks and workflows that relate to the production of letters. Letter production flows in two directions; incoming and outgoing. Incoming letters contain multiple

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

forms of implicit and explicit information. Processing this information results in tasks such as responding to requests, storing physical and digital copies of the letter, applying a medical coding scheme, and using information contained in the letter to inform or modify patient treatment and/or other clinical decisions. Outgoing letters involved multiple processes as well, such as audio dictation, transcription of audio text, application of preformatted templates, appending directions and notes to audio files or digital documents, referencing information in the electronic medical record, archiving letters, and posting the letter. The specific details of how this work is carried out varies from practice to practice. The workflow can look remarkably different in a single-site independent practice when compared to the workflow in a super-practice, sharing services across multiple sites. A single-site practice may allocate this work to receptionists who process letters, while a super-practice might outsource all medical coding, audio transcription, and typing to a central location.

The following workflow was observed at a small single site practice. Receptionists sorted incoming letters into groups, giving them red, amber, or green status dependent on response urgency. Then, they scanned the letters into the system and used a commercial software system called DocMan to sort them using the colour-coded system. One of the receptionists is assigned the task of “dealing” with the letters. This involves opening a letter that has been briefly coded by a GP with the highlighter tool in the document management software, reviewing the highlighted medical terms or phrases, and allocating appropriate Read codes. The letter is then archived on the system, the task is marked complete, and the receptionist continues to the next letter. If the receptionist has a question about the appropriate Read code, text comments are added to the document, which is then saved, and assigned as a task on the system, i.e. sent back to the GP. Other types of letters, for example, from a medical supplier requesting information about a patient who requires a specific device, must find the required information in the relevant health record, then use a template letter to respond to the supplier.

The length of letters ranges from precise and short to long and verbose. Letters have a different utility dependent on who is writing them, who is reading them, and the context. For example, during our research a GP was observed reading a 9-page letter from secondary care. The GP skim-read the letter on her monitor and after a few minutes declared, “ah hah, you see here it is”, highlighting a paragraph with the mouse cursor, saying “this is all I need, out of this whole document, this is what is relevant to me.” She continues: “I can tell the rest of this letter is for their [the specialist’s office] purposes, it’s a form of documentation, I don’t need all this.” The GP did not require most of the

content of the letter, only a single paragraph. Her prior experience of the specialist who wrote the letter enabled her to extract the relevant information efficiently and discount the rest.

Documentation of clinical consultations

A scenario in which a system could automate clinical documentation using natural language processing and text summarisation was described to staff in focus groups and interviews, provoking lengthy discussion. Some participants saw it as a potentially useful technology, with multiple “when can we have that” comments, while others felt it might have negative consequences on the clinical encounter and could even undermine the skills and critical thinking of the clinician. One GP argued that having a transcript of the entire conversation would be a waste of time, producing a great deal of unnecessary information. It is not uncommon for GPs to re-write or edit their notes, particularly notes for new cases or complex circumstances, or when the notes are written by trainees. What may look like a relatively simple administrative task in fact involves a considerable amount of clinical skill and experience. Removing this task from the clinical workflow would remove an opportunity for the clinician to reflect critically about an interaction. Research shows that writing engages the brain, allows GPs to be better observers, supports empathy, and engages critical thinking [34].

Scope for automation

The online survey of experts in automation systems achieved 156 responses from academic and industrial experts around the world, providing a total of 4,599 task level ratings. The distribution of respondents by country and qualification is shown in figures 2 and 3.

[Figure 2: Number of survey respondents by country.]

[Figure 3: Number of survey respondents by qualification.]

The Gaussian Process machine learning model that we applied to the data suggests that about 44% of the administrative tasks performed in primary care are “mostly” or “completely” automatable using current technology. Figure 4 shows the distribution of automatability for all administrative tasks performed by general practice staff.

[Figure 4: Distribution of Automatability Scores. Plot shows the Kernel Density Estimate of the Scores.]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

As discussed above, tasks performed by various occupations overlap, but the automatability of those tasks are independent of the occupation, i.e. an automatability score is not dependent on who performs the task. For example, medical coding has an estimated automatability of 3.2. Thus, automation of this task, which is mostly performed by receptionists, summarisers, and secretaries, but also on occasion by GPs depending on the practice, could either remove this task from their workloads altogether, if the systems are considered sufficiently reliable, or reduce it to checking codes produced by an automated system. Some of the most highly automatable tasks in our data include: *payroll and managing finances, checking and sorting post, printing letters, communicating with patients through texting, management of paper archives (onsite or offsite), transcription, email account management, letter scanning, checking for errors in paperwork, and internal communications (for example messages to staff or new employee inductions).*

To further understand the automatability predictions, three categories were generated: automatable (predicted score is >3.0), not-automatable (predicted score < 2.0), and partly-automatable (predicted score ≥2.0 to ≤3.0). Figure 5 presents the predicted automatability scores of administrative occupation tasks classified into these risk categories.

[Figure 5: Percentage of Administrative Tasks performed in Primary Care by Automatability Categories.]

Scanning clerks are typically, but not always, part time employees that are responsible for scanning physical documents (mostly letters) and creating digital objects (usually a PDF file) and attach these files to the patients’ health record or import files into other document management systems. Scanning clerks also add metadata to the PDF so the files are searchable and contain all contextual information. Results show that only the work of scanning clerks could be fully automated using currently available technology, thus eliminating the need for that occupation entirely. Many routine administrative tasks and typical office work (e.g. answering phones, writing letters and managing email) are undertaken by all occupational groups, but the impact of automation will vary according to the number of tasks in each automatable category that they perform. Figures 6 and 7 below show the potential impact of automation technologies by occupation for clinical and non-clinical staff.

[Figure 6: (KDE) Distribution of Task Automatability for Primary Care Occupations (non-clinical staff).]

[Figure 7: (KDE) Distribution of Task Automatability for Primary Care Occupations (clinical staff).]

Further detailed results can be found in the technical report for The Health Foundation, currently in preparation [35].

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Discussion

We believe this study is the first to combine detailed validated observations of workflows and a specially developed ordinal scale of automatability to assess the potential for automating various administrative tasks routinely carried out by primary care staff. Our findings indicate considerable scope for extending the use of automation in general practice. It is important to acknowledge that certain capabilities of automation already exist in primary care. In some practices scanning clerks put paper documents into scanners that turn physical text into digital type using optical character recognition. Another form of automation is the use of templates and pre-generated text in desktop word processors or in electronic medical record notes. Automated report generation and custom search queries to input patient data to different systems is another form of automation that is sometimes used. Our analysis has looked at the scope for further automation of tasks and we conclude that it is considerable.

A few caveats must be borne in mind. Observations took place in only six general practices. These varied in terms of list size and geographical location, but we cannot claim that these were representative of all practices in England. The survey of automation experts was, we believe, the largest of its kind to date and it provides a robust baseline estimate of the extent to which tasks are automatable using only currently available technology, but participants did not have specific expertise in the application of automated systems to healthcare. We made up for this deficiency by inviting discussion and feedback from primary care staff involved in interviews and focus groups, but they, in turn, lacked direct experience of using the automated systems that might eventually modify their working practices.

Automation technologies will be most useful when they provide clinicians with the right information at the right time; employing filters of context and urgency to remove details that are irrelevant or distracting. Automation of some mundane tasks, such as telephone answering, letter writing, document scanning, email monitoring, and information filtering, could free up valuable time for patient-facing tasks. . We find that many tasks involving data entry, sending information, formatting information, maintaining records, transcription, and operating office equipment to name a few, have high potential for automation. Many practices have basic forms of automated telephone answering, but more sophisticated systems are now available that would enable a voice assistant to help triage the hundreds of phone calls that a practice can receive in a day. However, there are drawbacks to automated phone systems that would need careful consideration and evaluation. Receptionists help prevent unneeded appointments by screening phone calls and getting patients to the right area for their appointment, these skills rely on social perception and an understanding of the patient's

history. Automated phone systems must be aware of this function and help guide patients through their case without scheduling unneeded appointments.

Clinical documentation could benefit from automation for three main reasons. First, these are tasks that every clinician must complete with a degree of urgency. It is considered good medical practice to document the patient consultation as soon as possible, i.e. it cannot be left to the end of the day. Second, documentation requires attention to detail and is time intensive. Even very brief encounters require accurate documentation. Third, this work is typically squeezed into a busy schedule, usually performed during the time it takes a patient to leave the examination room and the next patient arriving. Vendors of electronic health record systems have advocated the use of automation to facilitate record-keeping by use of structured data formats [36–38]. Natural language processing can be used to dissect patient-doctor conversations and create notes [39]. How to ensure the validity and reliability of such systems has been the subject of much research and remains a controversial topic [18,40]. However, a recent survey of UK GPs revealed considerable optimism about the likelihood that clinical documentation would be fully automated in the near future, with a large majority believing this would occur within ten years [41].

Documentation of clinical consultations is a good example of why automation should be approached with caution. Automating this type of work could have unintended consequences. Documentation of clinical encounters serves as a “tool for thinking” for doctors and as an important means of coordination and communication [40]. Reading and writing clinical notes engages the medical professionals’ critical faculties in a useful way. Allowing machines to take over would circumvent the way clinicians have been trained to document clinical encounters [42]. It would be important to identify the skills medical professionals would need in order to successfully work with automation technologies while maintaining a high level of care quality.

Automation technologies can be expensive and time-consuming to implement, so those contemplating the purchase of such systems should consider the following issues: Which types of tasks would be most beneficial to the practice if automated? Are those tasks amenable to automation? If this results in a change in occupational roles, who will it affect most? How will staff be trained and supported to use the new technology? If tasks are removed from the workload of a staff member, what kind of work (if any) will replace the task? And who will be responsible for the maintenance of the tasks while new workflows are in development?

Some tasks with high automatability scores will be challenging to implement for social and organisational reasons. Previous research has shown that staff contribute to patient quality and safety through hidden ways that emerge in the moment and require tacit knowledge [43]. The tacit

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

and contextual work, such as reading social cues, currently lies outside the capabilities of automated systems and machine learning technologies [29]. Healthcare is crucially dependent on these human skills and patient care would suffer if automation is pursued to their detriment. Our observations revealed a type of work that is often driven by exceptions, interruptions, competing critical requests and time sensitivity. Work practices that may appear routine, inefficient or superfluous, often have an intrinsic value for cooperative teamwork [44]. While our quantitative findings reveal considerable potential for automation, our in-depth fieldwork demonstrated the need for a cautious approach given the complexity of social systems in healthcare delivery environments.

Observing and gathering data on tasks in primary care presents at least three major challenges: task variance, task fit, and clinical or administrative task categorization.

First, like primary care itself, tasks are greatly varied. Variance in tasks can occur in the order in which parts of the task are performed, how long they take, the occupational role of the person performing the task, the “importance” of the task or how time-critical it is, and how many individuals become involved in completing it.

There are further complexities that emerge due to the task being performed in a healthcare context, when attention is given to patients' needs, backgrounds, and particular problems. In this context many administrative tasks require specialist medical knowledge. This is what makes work in healthcare different and exceptional, when compared to other fields with similar task descriptions.

The second challenge is that the same occupation does not always perform the same set of tasks across practices. We found that some tasks that receptionists routinely perform in one practice were performed by administrators at another practice. The practices we observed differed considerably in their organisational forms, from single-site practices working independently to larger groups or “super-practices”, who share services across multiple sites. The issue of “task fit”, i.e. allocating tasks to the appropriate person, involves both matching the task to the most appropriate pay grade and making sure clinicians are shielded from unnecessary administrative work. The driving force behind the assigning of tasks to occupations in this study was to accurately represent the core identity and “scope of work” of each observed occupation.

The third issue is the fundamental difference between “clinical” and “administrative” tasks. There is an abundance of administrative and clerical tasks that occur in primary care. While the field researcher observed, catalogued, and categorized nearly every clinical interaction that takes place in the domain of primary care, it was often hard to distinguish administrative tasks from those that are

purely clinical. Also, the O*NET database that was used for our analysis provided more terminology and language to describe administrative tasks and office work than it did for nuanced clinical work.

Conclusion

Evidence from our study show there is great potential to impact workloads in primary care by automating certain administrative tasks (approximately 44%, broadly). We believe that careful introduction of currently available technologies has considerable potential to reform administrative workflows and increase productivity in primary care. However, this should not be seen as a magic bullet that will solve the workload crisis in the immediate future. Introduction of new forms of automation should proceed cautiously with ongoing evaluation of their wider impacts.

Competing interests

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: all authors had financial support from The Health Foundation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Author Contributions

Matthew Willis: Writing, figure 1, coordination of authors, literature search, citation management, study design, formatting, draft iteration, data collection, data interpretation, manuscript submission, revisions.

Paul Duckworth: Figures, draft writing, editing, data collection, data analysis, literature search, data interpretation, revisions.

Angela Coulter: Reviewing and editing, writing, data interpretation, study design, revisions.

Eric T. Meyer: Reviewing and editing, writing, data interpretation, study design, revisions.

Michael Osborne: Data analysis, data interpretation, technical advising, figures, data collection, study design.

Funding

This work was supported by The Health Foundation, Award #7559. The study funders played no role in the study design, writing of the report, or the decision to submit the manuscript for publication. Researchers were independent of influence from study funders.

Data Sharing

All data relevant to the study are included in the article.

References

1 Hopson C. The State of the NHS provider sector. 2016. <https://www.nhsproviders.org/news-blogs/news/nhs-trust-leaders-warn-staff-shortages-now-outweigh-fears-over-funding>

2 Baird B, Charles A, Honeyman M, *et al*. Understanding pressures in general practice. London, UK: 2016. <https://www.kingsfund.org.uk/publications/pressures-in-general-practice>

3 Roland M, Everington S. Tackling the crisis in general practice. *BMJ* 2016;**352**:i942. doi:10.1136/bmj.i942

4 General Practice Forward View. NHS Engl. 2016.<https://www.england.nhs.uk/gp/gpfv/about/>

5 Clay H, Stern R. Making Time in General Practice. Prim. Care Found. Alliance. 2015;;1–83.

6 Byrne L, Bottomley J, Turk A. British Medical Association Survey of GPs in England. London: 2016.

7 Gibson J, Checkland K, Coleman A, *et al*. Eighth National GP Worklife Survey. Manchester: 2016. <http://research.bmh.manchester.ac.uk/healthconomics/research/Reports/EighthNationalGPWorklifeSurveyreport/>

8 Hobbs FDR, Bankhead C, Mukhtar T, *et al*. Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007–14. *Lancet (London, England)* 2016;**387**:2323–30. doi:10.1016/S0140-6736(16)00620-6

9 Hitchcock ALK, Sundorff E. Work in Progress: Towards a leaner, smarter public-sector workforce. 2017.

10 NHS statistics, facts and figures. London, UK: 2017. <https://www.nhsconfed.org/resources/key-statistics-on-the-nhs>

11 Fit for the Future: a Vision for General Practice. London, UK: 2019. <https://www.rcgp.org.uk/policy/future-vision.aspx>

12 Frey CB, Osborne MA. The future of employment: How susceptible are jobs to computerisation? *Technol Forecast Soc Change* 2017;**114**:254–80. doi:10.1016/j.techfore.2016.08.019

13 National Center for O*NET Development. 2016.https://www.doleta.gov/programs/onet/eta_default.cfm

14 Jeremy Bowles. The computerisation of european jobs. Bruegel. 2014.<http://bruegel.org/2014/07/the-computerisation-of-european-jobs/>

15 Pajarinen M, Rouvinen P. Computerization Threatens One Third of Finnish Employment. Published Online First: 13 January 2014.<https://econpapers.repec.org/paper/rifbriefs/22.htm> (accessed 27 May 2019).

16 Knowles-Cutler A, Frey C, Osborne M. Agile Town: The Relentless March of Technology and London’s Response. London, UK: 2014. <https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/uk-futures/london-futures-agiletown.pdf>

17 Jamieson GA, Skraaning G. Levels of Automation in Human Factors Models for Automation Design: Why We Might Consider Throwing the Baby Out With the Bathwater. *J Cogn Eng Decis Mak* 2018;**12**:42–9. doi:10.1177/1555343417732856

- 18 Kreimeyer K, Foster M, Pandey A, *et al.* Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review. *J Biomed Inform* 2017;**73**:14–29. doi:10.1016/j.jbi.2017.07.012
- 19 Ng A. What Artificial Intelligence Can and Cant Do Right Now. *Harv Bus Rev* <https://hbr.org/2016/11/%0Awhat-artificial-intelligence-can-and-cant-do-right-now>
- 20 Bresnahan T, Brynjolfsson E, Hitt L. Information Technology, Workplace Organization and the Demand for Skilled Labor: Firm-Level Evidence. Cambridge, MA: 1999. doi:10.3386/w7136
- 21 Brynjolfsson E, McAfee A. *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy.* 2012.
- 22 Brynjolfsson E, McAfee A. *The second machine age : work, progress, and prosperity in a time of brilliant technologies.* W. W. Norton & Company 2014.
- 23 Duckworth P, Graham L, Osborne M. Inferring Work Task Automatability from AI Expert Evidence. In: *ACM conference on Artificial Intelligence, Ethics, and Society.* Honolulu, Hawaii: 2019. <http://www.aies-conference.com/accepted-papers/>
- 24 Susskind RE, Susskind D. *The future of the professions: how technology will transform the work of human experts.* OUP Oxford 2017.
- 25 Acemoglu D, Restrepo P. Robots and Jobs: Evidence From Us Labor Markets. *Nber Work Pap Ser* 2017.
- 26 Arntz M, Gregory T, Zierahn U. Revisiting the risk of automation. *Econ Lett* 2017;**159**:157–60. doi:10.1016/j.econlet.2017.07.001
- 27 James Manyika, Michael Chui, Mehdi Miremadi, *et al.* Harnessing automation for a future that works. McKinsey Glob. Inst. 2017. <http://www.mckinsey.com/global-themes/digital-disruption/harnessing-automation-for-a-future-that-works>
- 28 Autor DH, Handel MJ. Putting Tasks to the Test : Human Capital, Job Tasks, and Wages. *J Labor Econ* 2013;**31**:59–96. doi:10.1086/669332
- 29 Bakhshi H, Downing JM, Osborne MA, *et al.* The Future of Skills: Employment in 2030. 2017.
- 30 Goos M, Manning A, Salomons A. Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *Am Econ Rev* 2014;**104**:2509–26. doi:10.1257/aer.104.8.2509
- 31 Kim H-C, Ghahramani Z. Bayesian Classifier Combination. In: *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics.* PMLR 2012. 619–627. <http://proceedings.mlr.press/v22/kim12.html>
- 32 Simpson E, Roberts S, Psorakis I, *et al.* Dynamic Bayesian Combination of Multiple Imperfect Classifiers. Published Online First: 8 June 2012. <http://arxiv.org/abs/1206.1831>
- 33 Willis M, Duckworth P, Coulter A, *et al.* The Future of Health Care: Protocol for Measuring the Potential of Task Automation Grounded in the National Health Service Primary Care System. *JMIR Res Protoc* 2019;**8**:e11232. doi:10.2196/11232
- 34 Wald HS, Borkan JM, Taylor JS, *et al.* Fostering and Evaluating Reflective Capacity in Medical Education: Developing the REFLECT Rubric for Assessing Reflective Writing. *Acad Med* 2012;**87**:41–50. doi:10.1097/ACM.0b013e31823b55fa
- 35 Willis M, Duckworth P, Graham L, *et al.* Future of Healthcare: Computerisation, Automation,

and General Practice Services. London, UK: 2019.

36 Rule A, Rick S, Chiu M, *et al.* Validating free-text order entry for a note-centric EHR. *AMIA . Annu Symp proceedings AMIA Symp* 2015;**2015**:1103–10.<http://www.ncbi.nlm.nih.gov/pubmed/26958249> (accessed 9 Sep 2018).

37 Lin SY, Shanafelt TD, Asch SM. Reimagining Clinical Documentation With Artificial Intelligence. *Mayo Clin Proc* 2018;**93**:563–5. doi:10.1016/j.mayocp.2018.02.016

38 Verghese A, Shah NH, Harrington RA. What This Computer Needs Is a Physician: Humanism and Artificial Intelligence. *JAMA* 2018;**319**:19. doi:10.1001/jama.2017.19198

39 Klann JG, Szolovits P. An intelligent listening framework for capturing encounter notes from a doctor-patient dialog. *BMC Med Inform Decis Mak* 2009;**9 Suppl 1**:S3. doi:10.1186/1472-6947-9-S1-S3

40 Bansler JP, Havn EC, Schmidt K, *et al.* Cooperative Epistemic Work in Medical Practice: An Analysis of Physicians’ Clinical Notes. *Comput Support Coop Work* 2016;**25**:503–46. doi:10.1007/s10606-016-9261-x

41 Blease C, Bernstein MH, Gaab J, *et al.* Computerization and the future of primary care: A survey of general practitioners in the UK. *PLoS* 2018;**13**:1–12. doi:10.1371/journal.pone.0207418 December

42 Willis M, Jarrahi MH. Automating Documentation: A critical perspective into the role of artificial intelligence in clinical documentation. In: *iConference*. Washington, D.C.: : Springer 2019.

43 Swinglehurst D, Greenhalgh T, Russell J, *et al.* Receptionist input to quality and safety in repeat prescribing in UK general practice: ethnographic case study. *BMJ* 2011;**343**:d6788. doi:10.1136/bmj.d6788

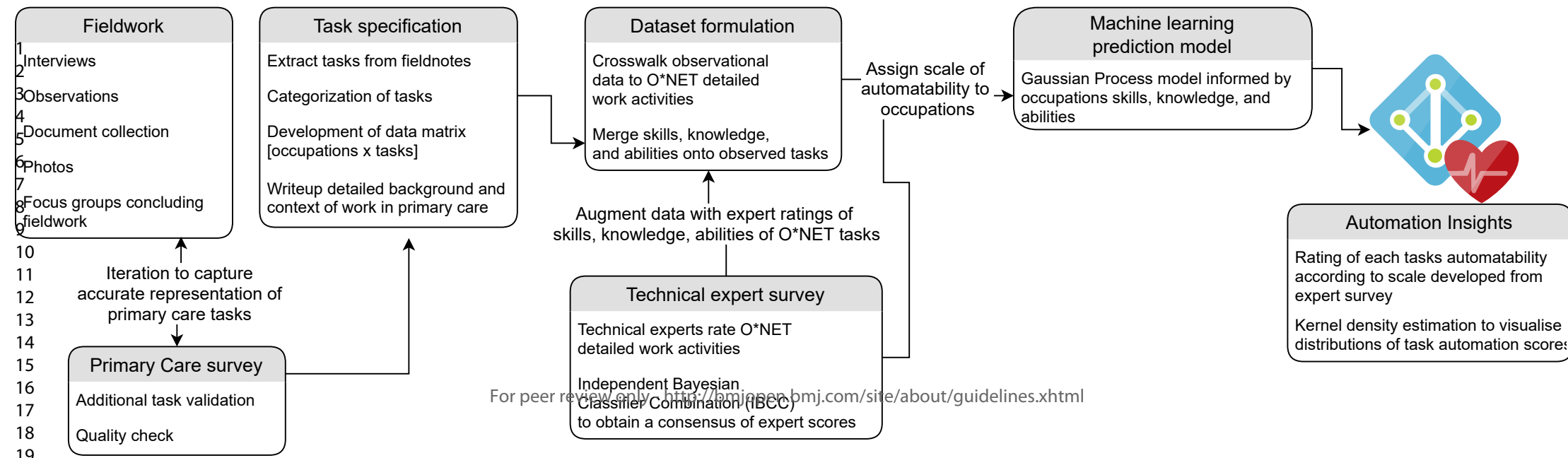
44 Cabitza F, Ellingsen G, Locoro A, *et al.* Repetita iuvant : Exploring and Supporting Redundancy in Hospital Practices. *Comput Support Coop Work* 2018.

Appendix

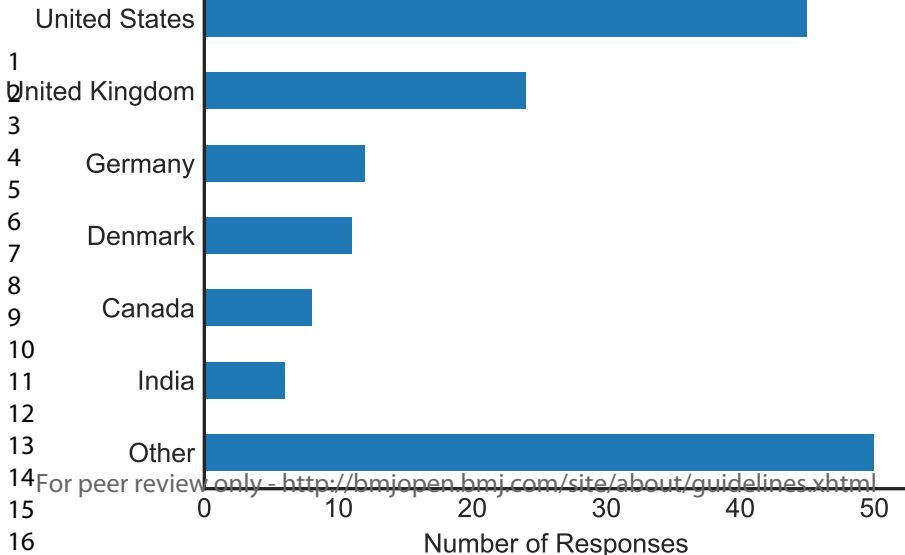
[Appendix Image 1: General Practitioner task list from O*NET database displayed in web portal report.]

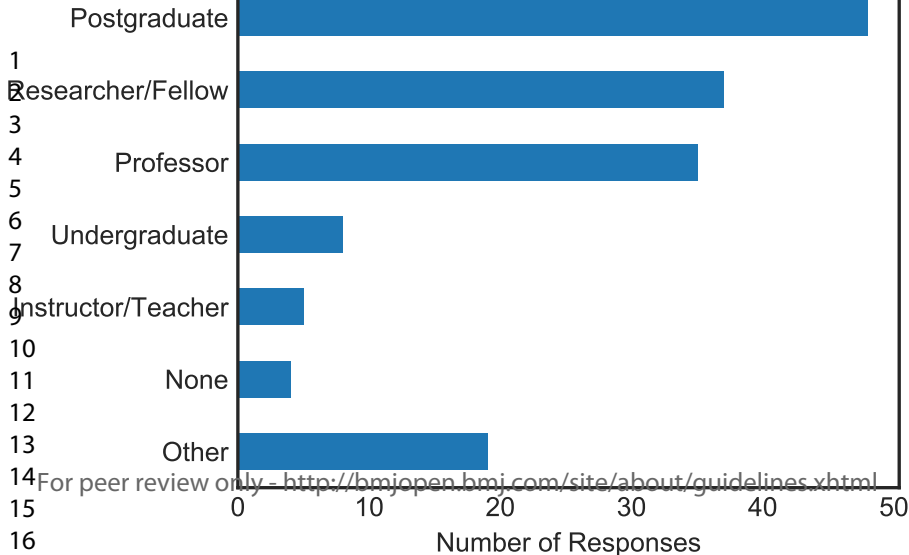
[Appendix Image 2: General Practitioner top knowledge and skills from O*NET database displayed in web portal report.]

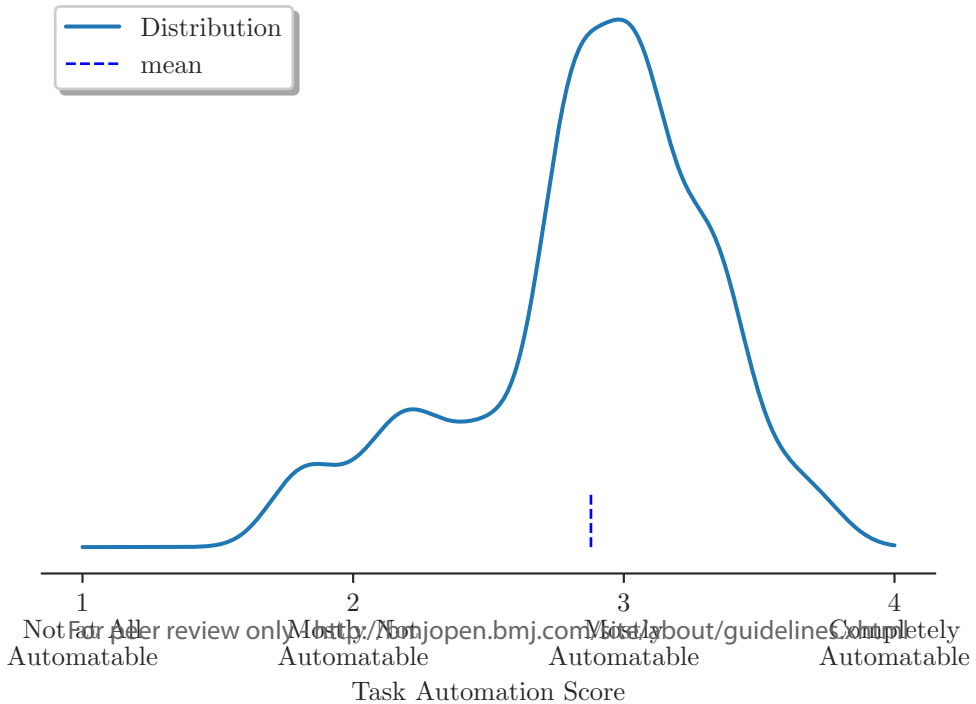
[Appendix Image 3: General Practitioner top abilities and detailed work activities from O*NET database displayed in web portal report.]



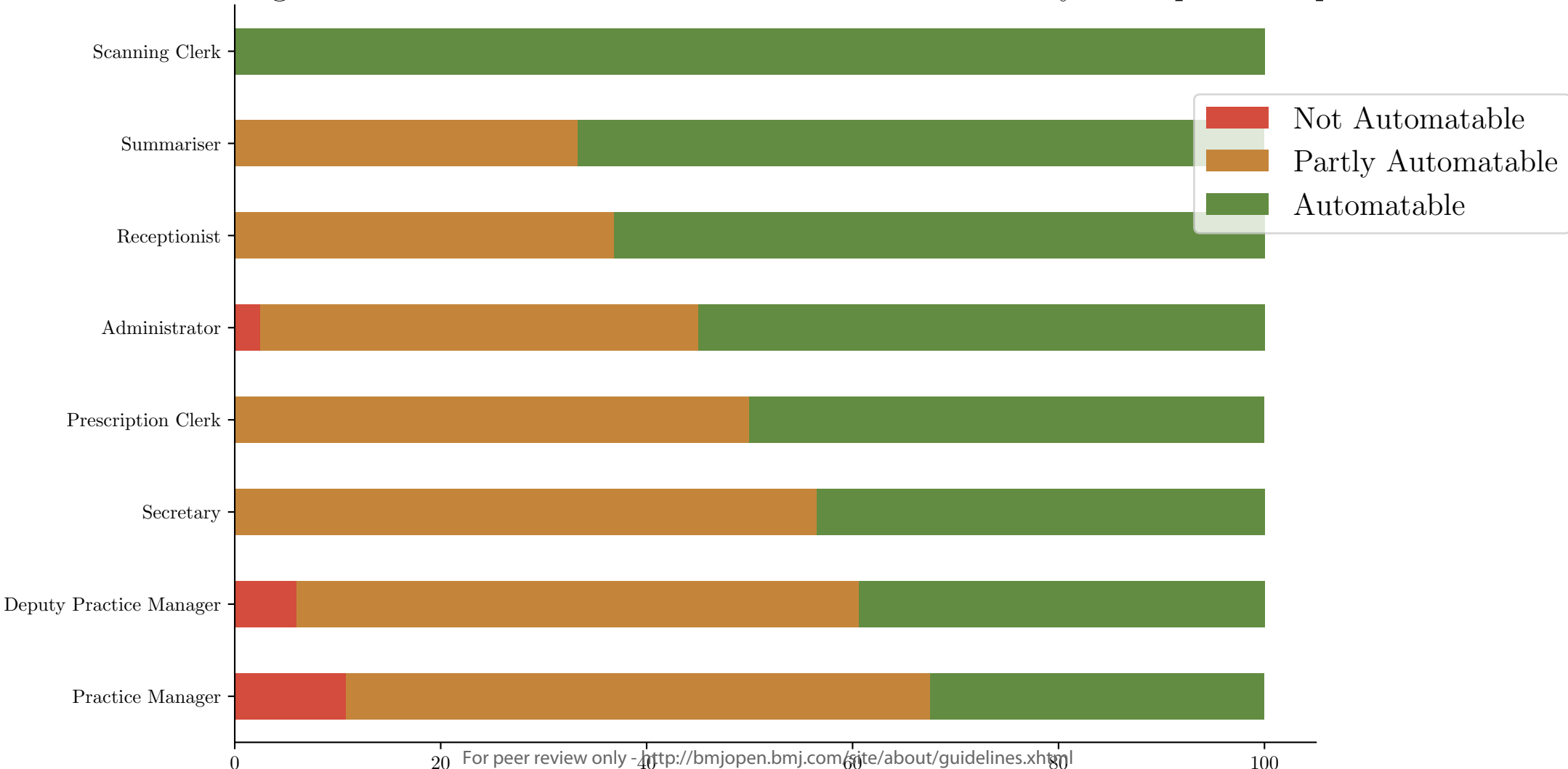
Top Responder Locations



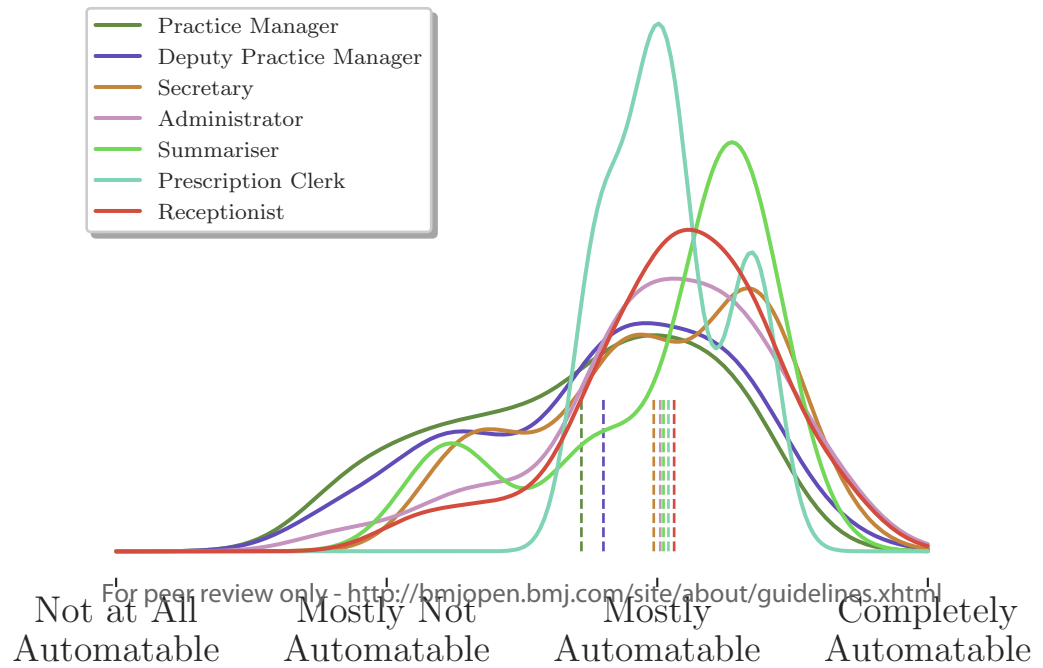




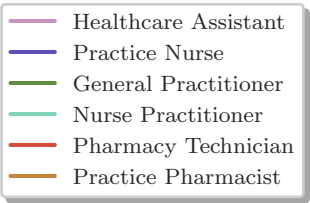
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34



Percentage of Tasks Performed by Automatable Category



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20



Not at All Automatable Mostly Not Automatable Mostly Automatable Completely Automatable

For peer review only - <http://bmjopen.bmj.com/site/about/guidelines.xhtml>

Physicians who diagnose, treat, and help prevent diseases and injuries that commonly occur in the general population. May refer patients to medical specialists when needed for further diagnosis or treatment.

Sample of reported job titles: Board Certified Family Physician, Family Medicine Physician, Family Physician, Family Practice Medical Doctor (FP MD), Family Practice Physician, Family Practitioner, Medical Doctor (MD), Medical Staff Physician, Physician, Primary Care Physician

View report: [Summary](#) [Details](#) [Custom](#)






[Knowledge](#) | [Skills](#) | [Abilities](#) | [Detailed Work Activities](#)

Save Table ([XLS](#)/[CSV](#))

All 10 displayed

Importance	Category	Task
95	Core	⊕ Prescribe or administer treatment, therapy, medication, vaccination, and other specialized medical care to treat or prevent illness, disease, or injury.
93	Core	⊕ Order, perform, and interpret tests and analyze records, reports, and examination information to diagnose patients' condition.
93	Core	⊕ Collect, record, and maintain patient information, such as medical history, reports, or examination results.
92	Core	⊕ Monitor patients' conditions and progress and reevaluate treatments as necessary.
91	Core	⊕ Explain procedures and discuss test results or prescribed treatments with patients.
88	Core	⊕ Advise patients and community members concerning diet, activity, hygiene, and disease prevention.
85	Core	⊕ Refer patients to medical specialists or other practitioners when necessary.
77	Core	⊕ Coordinate work with nurses, social workers, rehabilitation therapists, pharmacists, psychologists, and other health care providers.
75	Core	⊕ Direct and coordinate activities of nurses, students, assistants, specialists, therapists, and other medical staff.







Importance Knowledge

100		⊕ Medicine and Dentistry — Knowledge of the information and techniques needed to diagnose and treat human injuries, diseases, and deformities. This includes symptoms, treatment alternatives, drug properties and interactions, and preventive health-care measures.
91		⊕ Psychology — Knowledge of human behavior and performance; individual differences in ability, personality, and interests; learning and motivation; psychological research methods; and the assessment and treatment of behavioral and affective disorders.
86		⊕ Customer and Personal Service — Knowledge of principles and processes for providing customer and personal services. This includes customer needs assessment, meeting quality standards for services, and evaluation of customer satisfaction.
85		⊕ Biology — Knowledge of plant and animal organisms, their tissues, cells, functions, interdependencies, and interactions with each other and the environment.
83		⊕ Therapy and Counseling — Knowledge of principles, methods, and procedures for diagnosis, treatment, and rehabilitation of physical and mental dysfunctions, and for career counseling and guidance.

[back to top](#)
Skills [Save Table \(XLS/CSV\)](#)

10 of 11 displayed

Importance Skill

85		⊕ Critical Thinking — Using logic and reasoning to identify the strengths and weaknesses of alternative solutions, conclusions or approaches to problems.
81		⊕ Active Listening — Giving full attention to what other people are saying, taking time to understand the points being made, asking questions as appropriate, and not interrupting at inappropriate times.
81		⊕ Judgment and Decision Making — Considering the relative costs and benefits of potential actions to choose the most appropriate one.
78		⊕ Reading Comprehension — Understanding written sentences and paragraphs in work related documents.
78		⊕ Writing — Communicating effectively in writing as appropriate for the needs of the audience.
75		⊕ Complex Problem Solving — Identifying complex problems and reviewing related information to develop and evaluate options and implement solutions.

	Importance	Ability
100		⊕ Problem Sensitivity — The ability to tell when something is wrong or is likely to go wrong. It does not involve solving the problem, only recognizing there is a problem.
94		⊕ Oral Expression — The ability to communicate information and ideas in speaking so others will understand.
91		⊕ Oral Comprehension — The ability to listen to and understand information and ideas presented through spoken words and sentences.
81		⊕ Inductive Reasoning — The ability to combine pieces of information to form general rules or conclusions (includes finding a relationship among seemingly unrelated events).
81		⊕ Written Comprehension — The ability to read and understand information and ideas presented in writing.
78		⊕ Deductive Reasoning — The ability to apply general rules to specific problems to produce answers that make sense.
75		⊕ Written Expression — The ability to communicate information and ideas in writing so others will understand.
72		⊕ Speech Clarity — The ability to speak clearly so others can understand you.

[Back to top](#)

Detailed Work Activities

Save Table (XLS/CSV)

10 of 21 displayed

- ⊕ Immunize patients.
- ⊕ Prescribe medications.
- ⊕ Prescribe treatments or therapies.
- ⊕ Analyze test data or images to inform diagnosis or treatment.
- ⊕ Collect medical information from patients, family members, or other medical professionals.
- ⊕ Order medical diagnostic or clinical tests.
- ⊕ Record patient medical histories.
- ⊕ Monitor patient progress or responses to treatments.
- ⊕ Explain medical procedures or test results to patients or family members.
- ⊕ Advise communities or institutions regarding health or safety issues.

Standards for Reporting Qualitative Research (SRQR)*

<http://www.equator-network.org/reporting-guidelines/srqr/>

Authors note: this study was a mixed methods interdisciplinary project. The results reported in this article are the quantitative findings that used qualitative ethnographic fieldwork to fine tune the quantitative model and ground it in actual observations. As such, some items on this checklist may not apply. The detailed qualitative findings, coding, analysis, and theorization are a separate project and will be reported in a follow up article.

Page/line no(s).

Title and abstract

Title - Concise description of the nature and topic of the study Identifying the study as qualitative or indicating the approach (e.g., ethnography, grounded theory) or data collection methods (e.g., interview, focus group) is recommended	1
Abstract - Summary of key elements of the study using the abstract format of the intended publication; typically includes background, purpose, methods, results, and conclusions	2

Introduction

Problem formulation - Description and significance of the problem/phenomenon studied; review of relevant theory and empirical work; problem statement	3 and 4
Purpose or research question - Purpose of the study and specific objectives or questions	4

Methods

Qualitative approach and research paradigm - Qualitative approach (e.g., ethnography, grounded theory, case study, phenomenology, narrative research) and guiding theory if appropriate; identifying the research paradigm (e.g., postpositivist, constructivist/ interpretivist) is also recommended; rationale**	4 and 5
Researcher characteristics and reflexivity - Researchers' characteristics that may influence the research, including personal attributes, qualifications/experience, relationship with participants, assumptions, and/or presuppositions; potential or actual interaction between researchers' characteristics and the research questions, approach, methods, results, and/or transferability	5
Context - Setting/site and salient contextual factors; rationale**	9
Sampling strategy - How and why research participants, documents, or events were selected; criteria for deciding when no further sampling was necessary (e.g., sampling saturation); rationale**	5 and 6
Ethical issues pertaining to human subjects - Documentation of approval by an appropriate ethics review board and participant consent, or explanation for lack thereof; other confidentiality and data security issues	5

Data collection methods - Types of data collected; details of data collection procedures including (as appropriate) start and stop dates of data collection and analysis, iterative process, triangulation of sources/methods, and modification of procedures in response to evolving study findings; rationale**	6 and 9
Data collection instruments and technologies - Description of instruments (e.g., interview guides, questionnaires) and devices (e.g., audio recorders) used for data collection; if/how the instrument(s) changed over the course of the study	6 and 7
Units of study - Number and relevant characteristics of participants, documents, or events included in the study; level of participation (could be reported in results)	9
Data processing - Methods for processing data prior to and during analysis, including transcription, data entry, data management and security, verification of data integrity, data coding, and anonymization/de-identification of excerpts	6
Data analysis - Process by which inferences, themes, etc., were identified and developed, including the researchers involved in data analysis; usually references a specific paradigm or approach; rationale**	6 and 7
Techniques to enhance trustworthiness - Techniques to enhance trustworthiness and credibility of data analysis (e.g., member checking, audit trail, triangulation); rationale**	6

Results/findings

Synthesis and interpretation - Main findings (e.g., interpretations, inferences, and themes); might include development of a theory or model, or integration with prior research or theory	9-11
Links to empirical data - Evidence (e.g., quotes, field notes, text excerpts, photographs) to substantiate analytic findings	9-11

Discussion

Integration with prior work, implications, transferability, and contribution(s) to the field - Short summary of main findings; explanation of how findings and conclusions connect to, support, elaborate on, or challenge conclusions of earlier scholarship; discussion of scope of application/generalizability; identification of unique contribution(s) to scholarship in a discipline or field	14-17
Limitations - Trustworthiness and limitations of findings	2 and 16

Other

Conflicts of interest - Potential sources of influence or perceived influence on study conduct and conclusions; how these were managed	17
Funding - Sources of funding and other support; role of funders in data collection, interpretation, and reporting	17

*The authors created the SRQR by searching the literature to identify guidelines, reporting standards, and critical appraisal criteria for qualitative research; reviewing the reference lists of retrieved sources; and contacting experts to gain feedback. The SRQR aims to improve the transparency of all aspects of qualitative research by providing clear standards for reporting qualitative research.

**The rationale should briefly discuss the justification for choosing that theory, approach, method, or technique rather than other options available, the assumptions and limitations implicit in those choices, and how those choices influence study conclusions and transferability. As appropriate, the rationale for several items might be discussed together.

Reference:

O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. **Standards for reporting qualitative research: a synthesis of recommendations.** *Academic Medicine*, Vol. 89, No. 9 / Sept 2014
DOI: [10.1097/ACM.0000000000000388](https://doi.org/10.1097/ACM.0000000000000388)